THE UNIVERSITY OF ALBERTA

AN INVESTIGATION OF FOUR METHODS OF PRESENTING PROGRAMMED MATERIAL

by

WALDEMAR R. UNRUH

A THESIS

PRESENTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF EDUCATION

DIVISION OF EDUCATIONAL PSYCHOLOGY

EDMONTON, ALBERTA

SEPTEMBER, 1962

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "An Investigation of Four Methods of Presenting Programmed Material" submitted by Waldemar Renaldo Unruh in partial fulfillment of the requirements for the degree of Master of Education.

ABSTRACT

The basic assumption underlying the study reported in this thesis is that it should be possible to develop a technology of instruction, broadly applicable to many areas of study; and that the field of programed instruction might be a fruitful one in which to search for the basis of such a technology. Before this technology can be developed, however, it seems necessary to examine a number of instructional methods and to determine which methods are most effective. Furthermore, various instructional methods may be more effective than others for certain types of pupils; and, about this too, more information is needed. The writer therefore undertook to compare four methods of presenting programed material to pupils at two different IQ levels. By this procedure it was felt that information would be obtained that would indicate which method of presentation was superior in general, and which method of presentation might be more effective for pupils at either a high or low level of intelligence. The relative effectiveness of the four instructional methods was predicted on the basis of a number of variables such as response mode, amount of feedback, reinforcement and probability of error response.

Forty children--20 high IQ and 20 low IQ-- in a Grade VII language class were chosen as subjects on the basis of a language test prepared specifically for this experiment

and given prior to it. This test was also used as the criterion post-test. The twenty pupils in each IQ level were assigned randomly to four groups of five each, thus providing for 10 subjects in each of four experimental programs, the quiz, modified quiz, no-practice and test programs. The subjects worked through the program on each of three successive days, and were given the post-test on the fourth day. The mean scores were then submitted to an analysis of variance and covariance so that they might be tested for significance after differences in pre-test, due to randomization, had been taken into account.

The results of the analysis of the data clearly indicate that the quiz and modified quiz methods are superior to the test and no-practice methods. An analysis of the mean time taken to complete the programs indicated that the modified quiz program required much more time than did the other three methods. Since there was little difference in achievement between modified quiz and quiz methods, it would appear that for this type of material, at least, the quiz method of instruction is more efficient.

The analysis of covariance also indicated a significant difference in IQ levels and a significant interaction effect; however, because the difference in high and low IQ groups for the no-practice method was much less than expected, and the difference in the test method was much larger than expected, this interaction has no useful interpretation.

ACKNOWLEDGEMENTS

The writer wishes to express sincere appreciation to Dr. J. D. Ayers for supervising the study and for his invaluable criticisms and suggestions; to Dr. A. Harper, Dr. W. Rundquist and Dr. A. Godwin for their constructive criticism; to Dr. S. Hunka for his advice regarding statistical procedure; and to the principal and staff of Elmwood School for generously permitting the writer to carry out his experimental work in the school.

TABLE OF CONTENTS

CHAPTER		PAGE
VI	DISCUSSION	99
	Presentation Method	99
	Interaction	102
	Time	103
	Error Rate	105
VII	SUMMARY AND IMPLICATIONS	106
	Summary of the Study	107
	Implications for Programming	108
	Implications for Further Research .	109
BIBLIOGRAPHY		112
APPENDIX A	Summary Data Related to Pupils Used in Experiment	118
APPENDIX B	Sample of Criterion Test	122
APPENDIX C	Sample of Programs used in Experiment	131

LIST OF TABLES

CHAPTER I

THE GENERAL PROBLEM

Programmed instruction[a] and the device known as the "Teaching Machine" has recently become one of the major topics of discussion among educators. Manufacturers of machines, publishers of programs and the layman involved in administration of the school system have all become involved. The danger exists that the effects, both positive and negative, of programmed instruction will be released and applied without sufficient attention to careful study of well-planned research. Much evidence is available which can be used appropriately to test this advance in education. An examination of this evidence, and further research based on it, will help to define some of the variables related to programmed instruction and will also result in a clearer approach to programming, making for optimal efficiency in its use.

For some time animal psychologists have been making intensive efforts to analyze and describe the learning process. With the advent of programmed instruction it becomes possible to apply what these scholars have discovered about the learning process to human learning. It may now be possible, and even necessary, to develop a technology of instruction which will, as a discipline, embrace previously discovered but unintegrated principles

---

a See p. 58 for definition of "programmed instruction" and related terms.

related to learning. This would obviate the need of subscription to a specific school of learning theory or psychology. The only criterion for accepting an instructional technique would be a pragmatic test of its effectiveness. Future research must therefore be in the form of combining known principles in some useful way rather than analyzing unrelated concepts without at least some attempt at synthesis.

Studies related to programmed instruction have, for the most part, amply demonstrated that such a method can be successfully used for instruction in many subject areas ranging from Keislar's (1959) work with Grade I children learning molecular theory to the rather esoteric studies of Gagne (1961, 62) in guided discovery. Further studies are cited by Stolurow (1961) to indicate that programmed instruction can be useful over a wide range of age and ability differences. Since the effectiveness of this mode of instruction cannot therefore be generally questioned, it is not primarily the purpose of this thesis to compare the classroom teacher and his techniques with the programs now available. It is more important that a study of various types of programs be undertaken to determine which of these can be most effectively used in a classroom situation.

It is the purpose of this study to investigate two specific areas related to learning and programmed instruction: Under what conditions will optimum learning take

place? What type of program will best meet the needs of the learner? This study will therefore be concerned with the way in which children learn best, the type of material from which they learn best, and the type of practice which leads to best results - that is, we are concerned with subject factors, task factors and practice factors. Furthermore, this study is committed to an examination of the evidence concerning interaction among the foregoing factors. Teaching machine programs and the way in which task, subject and practice factors interact in the learning situation, are made the explicit object of study. The type of program presentation used is really determined by the way in which these factors are combined, and it is this possibility of combination which makes it possible to submit the various types of programs to experimental testing. It is proposed that several types of program presentations be used to test their relative effectiveness among various types of pupils. This should provide some basic knowledge related to the technology of instruction as distinct from the principles of learning and a theory of learning. Among other things, it should provide us with some specific knowledge of:

a) the optimum combination of factors known to be conducive to efficient learning.

b) the limits of complexity required of a programmed teaching device.

c) the minimum amount of overt activity required for

effective learning.

d) the relationship between the type of program used and the intelligence of the pupil.

In other words, it is the purpose of the study to examine the merits of several instructional methods and to try to determine at least some of the factors which make these methods successful.

CHAPTER II

THE RELATED LITERATURE

Most of the current research in programmed instruction is intended to answer one of four basic types of questions.

1. The first and oldest type of research literature deals with the problem of evaluating programmed instruction on the basis of appropriateness to specific subject areas and various levels of ability.
2. A second type of research deals with the problem of evaluating and revising a specific program.
3. A third type of question which has been investigated deals with basic problems such as response mode, pacing problems, comparison of one type of programming with another.
4. A fourth type of research deals with the more basic psychological problems such as interaction effects of IQ, age and other learner variables.

The review of the literature in this thesis will depart somewhat from the pattern suggested by the problems listed above as it is felt that the following pattern is more useful than an historical approach. Studies will be reviewed dealing with the automated instruction device as such. This will be followed by a discussion of learner variables which are closely related to programmed instruction. A section will then follow on variables relating to

programming and the review will conclude with a brief criticism of current research.

Teaching Machines

Functional Analysis As early as 1866 Halcyon Skinner developed a teaching aid in the form of a spelling machine. In 1873 a machine was patented which it was claimed could teach logic. H.B. English (1942) developed a device used in the first World War to teach soldiers to squeeze a rifle trigger. Pressey (1922) invented several devices to automate testing. Many machines and devices have been developed recently and some logical classification, based perhaps on the functions which they perform, may prove useful. The first step in an analysis, however, is to determine how a teaching machine differs from a teaching aid.

There appears to be some disagreement as to the essential characteristics which a teaching machine must possess, and very little definitive evidence is available which would clearly indicate just what the limits of sophistication might be.

Porter (1957) has suggested three essential features which distinguish between teaching machines and teaching aids:

1. A sequence of carefully graded problems leads the learner to meaningful knowledge. Thus the teaching machine is not just another gadget to aid the memory but encourages the student to think.

2. A teaching machine requires some action on the part of the learner. Such action may be overt or covert.

3. The device provides for immediate knowledge of results. Such knowledge of results may be in the form of a signal indicating whether the answer is right or wrong or it may be in the form of clues or guides helping the student to formulate the correct answer. In any case guidance is so given that the student is prevented from following some line of error.

Carr (1959) suggests that a self-instructional device has four major components:

1. a display by means of which the program is presented,

2. a response panel which the learner uses in framing or preparing his response,

3. a confirming mechanism or some type of comparator to which the learner can prepare his response, and

4. a reinforcement mechanism which keeps the student moving.

In contrast to the specifications suggested by Porter and Carr, other experimenters have found that simpler devices may do an equally effective job; but it must be submitted that many of the machines now in use are research devices. Cram (1961) suggests that a teaching device merely provides information, asks for periodic responses, waits for these responses and then informs the pupil

of the appropriateness of the response before allowing him to go on to the next frame. A survey of the literature in this area would suggest that there are three crucial characteristics present in an effective teaching machine: the machine must make provision for pacing by the pupil; it must provide for action on the part of the learner - either overt or covert; and it must provide knowledge of results.

Classification of Teaching Machines and Devices. Since teaching machines can differ in so many ways, some method must be presented for classifying them. Perhaps the most satisfactory classification to date is that of Stolurow (1961). He uses as a differentiating concept the adaptability of the device. In addition to this he classifies each device on the basis of a further series of variables which he describes as characteristics of a maximally adaptive machine. These ten characteristics are found to some extent in all of the devices known as maximally adaptive machines but may be entirely missing on some of the less adaptive ones. These characteristics will be discussed briefly and will be followed by Stolurow's analysis.

Presentation - This involves the display of the material which may be in the form of a film or printed material.

Response - This may be in the form of selecting an answer from multiple choice list, or it may be constructed by the pupil either mechanically or in some written form.

Pacing - Either the machine or the learner does the pacing. In the simpler devices this is left to the learner unless there is an outside source such as the teacher to do the pacing.

Comparator - This provides a means whereby the pupil can compare his response to the correct one. As the machine becomes more adaptive, it takes over more and more of the comparator function.

Feedback - A minimally adaptive machine usually does not provide feedback, in that it merely indicates whether the answer is right or wrong but does not indicate the extent of correctness or incorrectness.

Collator/Recorder - None of the minimally adaptive machines is able to collate information.

Selector - This also is present only in adaptive machines. In the simpler forms of equipment it is possible, however, for the teacher to select material suitable for the individual pupil.

Library - The library in each case may be basic only, or in the more adaptive machines the library may be varied.

Program - In the minimally adaptive machines the program is always linear. In the more sophisticated machines a branching program is usually available.

Computer - This is present only in the more expensive adaptive machines.

1. Minimally Adaptive Machines: In this case the devices may be said to be incomplete by themselves because at some stage in the learning process they need some form of auxilliary help, usually in the form of the experimenter or the teacher. An example of such a device might be a deck of flashcards with stimulus or cue material on one side and the associated response material on the other. The cards are presented in some predetermined order; the pupil responds to the stimulus and checks his answer by examining the other side of the card. Usually there is a predetermined interval of time between the presentation of the stimulus and the response check. This period of time is controlled by the experimenter or teacher. In any case, the learner

sees the right answer whether he has made the right response or not. This means, of course, that the program is identical for all learners both in content and in the speed at which it is presented. A more sophisticated type of device is the memory drum. Here the items are displayed mechanically. In its usual form the drum, though it may pace the learner automatically, usually requires that some person be present who can inform the pupil of the rightness or wrongness of the response. Since some source beyond the machine is necessary for efficient learning to take place, this may still be defined as machine-program-teacher-learner system. It is still minimally adaptive.

2. Partially Adaptive Machines: Essentially there are two types of partially adaptive machines - the hardware type and the printed type. Most of these devices are used to measure recognition and require a multiple-choice type of response. A few require that the learner compose his response through the use of keys, slider, etc., or by writing on a separate answer sheet which can later be compared to the correct response. Some of the machines perform the comparator function for the pupil, but in the write-in devices this is not yet technically feasible at any complex level of instruction. In most of the devices the learner is given direct feedback by being shown the correct answer once he has indicated his own choice.

Another characteristic of these devices is that they

are paced by the learner. In most cases the learner must give the correct answer before he can proceed to the next frame. Examples of such devices are Pressey's multiple choice machine and Skinner's disk machine and arithmetic machine. A more intricate machine is the Western Design Autotutor, described as an automatic, random-access recording, microfilm, motion picture projector.

> "The student finds the first unit of subject-matter information on image 1 of the microfilm, along with a multiple-choice question based on that unit of information. He then selects an answer to the multiple-choice question, enters into a keyboard the frame number accompanying the answer he selects -- a frame number is always located beside each alternative answer -- and presses the 'view' button. The device then automatically locates and projects the image that is at the address that was the student's choice. If the correct answer to the multiple-choice question was chosen, the image selected will not only contain knowledge of results but also will contain the next unit of information, and the next question. If the incorrect one of the alternatives is chosen, the corresponding image will tell the learner of his error, provide information designed to correct the particular error, and direct him to return to the image at which he made his mistake and try again." (Stolurow, 1961, p.31)

Another method of presenting material similar to the way in which it is presented by the Autotutor is to use a printed device known as a scrambled text. Variations of this form of presentation are also available -- for example, the punchboard, the programmed text, the cut-back page booklet and the tab-type page. In all cases information is given and a multiple-choice type question asked. The pupil makes a choice among the options available.

If he is correct, he is allowed to go on; if he is in error he is referred to additional material and then allowed to proceed to the main stream of frames either by returning to the point at which he made the error or by following through some branch. This makes some adjustment to individual differences possible and requires less work on the part of the teacher in the selection of materials.

3. Adaptive Machines: Adaptive machines are also found in two basic forms. The first is a modification of an IBM computer used by Rath, Anderson and Brainerd (1959) to teach binary arithmetic. It allows for changes in teaching method without the building of a whole new machine. The pupil views the problem and proceeds to enter his response on a keyboard. The computer is able to check each step in the solution immediately. If an error is made in any step the machine stops the pupil immediately so that a line of error is discontinued. If the learner is able to construct the correct answer, new material is presented; if he makes an error response, the machine can so adjust the program that weaknesses in knowledge can be corrected. The advantage of this type of machine is, of course, that it can present the learner with problems at a suitable difficulty level.

A second type of adaptive machine, and the most complicated one, is the one produced by Pask (1958). This machine is under organic control -- that is, the teaching

machine is so designed that it seeks a compromise with the learner. The computer type of machine discussed previously tends to impose a rather rigid behavior pattern on the student, but in the case of the Pask machine, the machine itself does the learning. It provides a stimulus; the pupil responds to this stimulus; the machine then adjusts the program to fit the needs of this particular pupil. In other words, the machine has the job of discovering the strategy which the pupil uses in solving a set of problems and then of programming information in such a way that the best possible progress can be made by the pupil. This system is often described as a game analogy.

<u>The Learner</u>

<u>Learner Variables</u> Most of the research on programmed instruction has been focused specifically on the program or on the learner rather than to possible interaction effects between variables related to both. Some effort is necessary to locate possible sources of interaction and to isolate some of the effects thus obtained. For example, there may be an interaction between IQ and the type of program used or between time taken to complete a program and its effect on scores when various types of programs are used. Many such relationships undoubtedly exist and should be investigated. In the following discussion, intelligence, motivation, memory span and age will be discussed as learner variables which may affect performance.

A number of research studies have made much of the apparent lack of correlation between ability test scores and gain scores in subject areas where programmed instruction has been used. That is to say, it appears that intelligence quotients may be fairly good indicators of final level of achievement, but correlations tend to be much lower between intelligence scores and gain scores, often approaching negative values. Such results have often led to the conclusion that duller pupils benefit more from programmed instruction than do brighter ones. For example, Porter (1959) found that in his experimental group of 6th graders there was a negative correlation between IQ and achievement.

Little (1934) used teaching machines to provide drill in educational psychology to college freshmen. His findings tended to indicate that slower students benefitted more from programmed instruction. He divided his classes into quartiles and matched two groups - a drill group and a control group. On the final test he noted that the lowest quartile of the drill group averaged 19 points higher in their scores than did the control group. The upper quartile only scored 5 points higher than the control group. He also found that the lowest quartile of the drill group was 7 points above the median for the control group.

The studies mentioned appear to support two possible hypotheses - namely, that there may be a relationship

between the type of programming and the ability of the pupil and that instruction by means of the teaching machine may result in greater homogeneity among the members of a class. A study by Briggs (1949) tends to support this line of thought. He found that when he used the punchboard device more pupils received A's and B's than before. The same trend was found in studies by Pressey (1958) and Jensen (1949). Whether this reduction in heterogeneity is due to more efficient teaching or to changes in motivation is not known. Individual differences in rate of learning may be due to a host of different variables and to interaction among these variables.

Motivation, or readiness to respond, is another important variable which may affect learning. It is also one of the most difficult to control experimentally. Many studies show remarkable increases in learning when programmed instruction is used; but they seldom control adequately for the increased motivation which accrues because of the presence of the hardware. Extensive long-term studies have not as yet been carried out to determine whether this added motivation is due to novelty or perhaps because of the intrinsic effect of machine upon the learner -- an effect different from that produced by a teacher or text book. Porter (1958) investigated the effects of year-long spelling instruction. He used Skinner-type machines to give instruction in spelling. He found that there was no

relationship between achievement and the number of lessons given and concluded that motivation does not decrease as time passes but that the number of tries does decrease. He noted too, that better pupils responded at a much faster rate, perhaps because of the release from regular classroom routines. Barlow (1960) suggests a further possibility and that is that progress will depend on how the pupil views the machine. For optimal effort he must see the machine as a self-teaching device and be prepared to take responsibility for his own instruction. This in itself may change his whole attitude toward learning and may account for some of the disparity between his present achievement and previous records. These previous records, even in the form of scores on standardized tests, may in effect have been scores indicating a present level of function rather than maximum capacity.

A further problem, not yet investigated, is one of finding optimum levels of motivation. Taylor and Spence (1952) have discovered an interaction effect which suggests that increased levels of motivation may be useful only for the learning of simple tasks but may very well hinder the learning of more complex processes. This would also support the theories of Berlyne (1960) regarding optimal arousal level. If substantiated, this finding would be relevant to teaching machines in that it would suggest that machines should be so constructed that they could change pace and

difficulty level to suit the needs of the pupil. Machines should also be constructed so that some rivalry, preferably against the pupil's own record, could take place. The result would be some level of ego-involvement and probably fewer errors.

Rate of learning may also be related to memory span and age. It may be that how well a pupil is able to learn a particular concept or specific body of knowledge depends upon his ability to recall past experiences related to the new learning. At earlier ages recognition memory will be of greater importance in programming. As the pupil gets older he will be able to remember items in either recall or constructed form. Furthermore, as Woodrow (1938) suggests, divergence in achievement tends to increase with practice under some conditions of subject matter and depending on past experience the amount of new information can be adjusted to be compatible with the known span for a particular type of material at a given age.

In spite of what has been said of the poor correlation between intelligence test scores and gain scores, it remains obvious that intelligence is a factor in the rate of learning. In programming for instance, it is important that we find the relationship between retention and intelligence so that reviews can be properly placed in the series of frames either as a part of a linear sequence or by the use of branches. This will help to take care of

individual differences in retention without the writing of programs for each individual. A further consideration is the problem of amount of reward and its relation to intelligence. Abel (1936) found that reward for improvement appeared to benefit the high-intelligence group most. He does not, however, take the type of reward into account, and this remains to be investigated.

A final note about gain scores may be appropriate here. The negative correlations found between intelligence and gain scores may be a form of statistical artifact. Since bright pupils usually know a good deal of the material in any subject area, it is difficult for them to make substantial gains. This is especially true where the criterion test does not have enough "top" to discriminate among brighter pupils. Slower pupils, especially underachievers, can, with the added motivation, increase their scores considerably. This may give rise to negative correlations which must therefore be carefully studied to avoid erroneous interpretations. In addition, it may be necessary to develop some new method of interpreting gain scores.

Programming and Special Groups of Learners When we examine learner variables, it is important that the research literature concerning the use of teaching machines and special groups of learners be reviewed. This will also involve a comparison of programmed instruction and conventional instruction.

Studies carried out with normal children have all been in agreement that programmed instruction can result in excellent learning and a saving of time. Osler (1960) used children aged 6 to 14 to teach concepts. She obtained a mean trial time of 15 minutes for 150 trials. Manual methods had been used previously and had required an average time of 65 minutes. Such a device could be used to teach reading readiness. In this case it would be necessary, however, to do further research because it would require that use be made of the pupil's present repertoire. It would then be possible to bring the appropriate verbal response under the control of non-verbal stimuli. Number concepts have been taught in a similar way. Alter, Eigen and King (1961) aimed to measure the efficiency of a preverbal program and "confirmation" and "confirmation and trinkets" reinforcement conditions with preliterate young children. The post-test indicated that the number of errors remained constant and there was a significant gain in learning which the authors rightly attribute to the fact that much of what was learned may have been in the form of learning to use the machine. Such misleading results could be avoided by giving children practice with the machine ahead of time.

Programming for Special Subject Areas The use of teaching machines in teaching spelling has been attempted in several studies. Perhaps the best controlled of these

is that done by Porter (1959). Second and sixth graders used Skinner-type machines to take instruction in spelling. A control group was given spelling instruction in the usual way while the experimental group used machines for 22 of the 34 weeks during which the course was given. In both grades the experimental group achieved significantly higher on standardized tests and used much less time to learn the material. It is to be noted, however, that what is attributed here to the teaching machine may, in fact, be the result of programming. If the regular instructor used the techniques employed by the programmer, the results might have been different. It will be necessary to isolate the factors that contribute to better learning in a specific area such as spelling.

In a 1960 study Meyer reported an experiment using programmed textbooks to teach Latin and Greek elements which have come into English as roots and prefixes. The experimental groups were chosen from superior grade eight readers. One group used programmed textbooks which had no answers in them but were corrected by the experimenter and returned later. A second group scored their own answers by putting X's on the incorrect ones, thus receiving immediate knowledge of results. This same procedure was used by an honors group. A fourth group marked the errors and also the pages on which they occurred so that they could return later to review this material. All groups,

with the exception of the honors group, were selected by means of a pre-test. The medians and ranges of the three groups on the test are reported as being approximately equal. The error score was significantly higher for those whose books were marked by the experimenter. Kennedy (1961) points out that this may be because those who scored their own work probably gained from the immediate knowledge of results or perhaps just from knowledge of results, whereas those whose books were scored by the experimenter may not have bothered to correct their errors. In any case the programmed textbook forces the student to know his results and is therefore a valuable form of instruction. Going back to review did not seem to be an improvement over just marking the errors. It was noted, however, that learning of material correlated well with error score and that keeping errors down may be important.

Some evidence is available on studies with groups of exceptional children. Kalin (1961) investigated the ability of superior pupils to learn advanced mathematics by means of programmed instruction. The subject matter was presented in a programmed text dealing with equations and inequalities and was given to ninety-five 5th and 6th graders whose IQ scores were all considered to be above 115. The pupils were randomly assigned to 10 groups, five of which acted as controls and five as experimental groups. Experimental subjects read the programmed text at their

own rate, without assistance. Control subjects were taught by regular teachers who had been prepared by means of conferences and a study of the program. The study was evaluated by means of a 44-item criterion test. Differences in test score between paired experimental and control were not statistically significant. A twenty per cent time saving for experimental groups was noted. Means for these 5th and 6th graders ranged from 26.2 to 30.8 compared with a 9th grade algebra class mean of 35.0. This latter result probably indicates that younger, bright children can be taught more advanced work, but, then, they can be taught almost anything, and one assumes that the grade 9 algebra class with which they are here compared was an average class. A study could be performed on duller pupils to see if they too can learn material which is considered to be too difficult for them.

As early as 1915 Ordahl and Ordahl had seen the possibilities of using machines to teach retarded children. Contrary to other findings, discussed elsewhere in this review, these authors found a positive correlation between gain and MA. Stolurow (1961) points out that this may be due to the heavy demand made upon the memory of these subjects. Perhaps there is a limited general level of ability which is needed and beyond this level memory is no longer revealed statistically as a factor in learning.

Evidence on the effectiveness of instruction by machine of retarded children is largely lacking, perhaps because of the difficulty of programming for such a group.

Other special groups of subjects have been submitted to automated instruction. Falconer (1959) attempted to teach deaf children by machine. His first studies were designed to teach word-recognition to young children. He programmed 15 nouns with 5-choice problems. Each of the nouns was combined with a picture and 6 different frames were prepared. The picture presented was the cue and the child was to match the cue to the word in the word list that was presented with the picture. Falconer used 8 deaf children in a beginning primary class. They used the machines for 5 minutes a day for 10 days. A 15-item pencil and paper test was administered both before and after the students took the machine instruction. The test results indicated that the children had learned approximately 14 words. This knowledge did not deteriorate with time as a retention test given two weeks later showed. Since no control group was used, however, it is difficult to make any comparison with conventional instruction.

Evidence on the effectiveness of machine instruction is rapidly accumulating in the field of adult education. The number of subject areas in which programs are now available to the public is great. Perhaps the most outstanding work of all in this field is that of Holland

(1959 a and b) at Harvard. The program was in the form of a programmed text based on Skinner's (1953) "Science and Human Behavior". The course was run for two years, being revised at the end of the first year on the basis of the weaknesses indicated in the course during the first presentation. During the first year 187 students took the course and during the second year 146 took it. These students were a mixed group consisting of sophomores, juniors, seniors, a few freshmen and some graduate students. The program was presented on disks in the Skinner disk machine. Each disk contained 29 frames. There were 48 such disks in the program the first year and this was extended to 60 disks in the second year in order to make the steps smaller, thus reducing the error rate. The median time taken to complete the course the first year was 14.5 hours and the second year 12.5 hours. It is of interest to note the range of time taken to complete the program. One student was able to complete the whole cycle of disks in 7 hours and 42 minutes while another took 26 hours and 20 minutes. Stolurow (1961) points out that this may indicate the presence of a great number of interacting variables difficult to control in a semester-long experiment.

Though many studies now report success in adult instruction by machine, there remains a good deal to be done in the identification of variables which affect the ability of adults to learn in a wide variety of subject areas.

These variables can best be interpreted in terms of the programs we prepare and it is to the evidence on program preparation that we now turn.

The Program

Rationale and Underlying Psychological Theory Early teaching devices, such as the ones used by Pressey, were really based on the theories of remembering and forgetting so prominent in the writings of behavioristic theorists of the 1920's. It was not until further study and research had broadened the scope of learning theory that it was realized that appropriate apparatus was needed toarrange for appropriate contingencies of reinforcement and a way of presenting a sequence of logical steps. Some means of comparing response given to response required was also needed if learning was to be more than a memory-drum type of activity. These were not the only factors which early experimenters had to take into account. Psychologists who had come under the influence of a number of schools of thought had by this time adduced a number of principles important to anyone attempting to develop a technology of instruction. For example, the S-R theorists emphasized pupil activity, the laws of frequency and effect, the place of practice to insure ability to generalize and discriminate and the place of conflicts and drive as a factor in human learning. The cognitive theorists, on the other hand, emphasized perceptual aspects of learning, learning with

understanding, feedback factors, goal setting and the arousal and reduction of cognitive dissonance. Finally, the personality theorists also contributed to our knowledge of the learner. They stressed individual differences as a result of physiological and social development, the problems of anxiety, motivation and group atmosphere. Any attempt to improve instruction must be based on some understanding of these principles.

Learning or the conditioning process may be defined as the acquisition of new responses or the strengthening of some response already in the subject's repertoire.

"We refer to the process of strengthening as <u>conditioning</u> and to its means of accomplishment as <u>reinforcement</u>". (Green, 1962, p.40) Reinforcement, in turn, is the result of a physical event acting upon an organism resulting in some change in its behavioral pattern. The Skinnerian theories of conditioning are based on this modified law of effect and will be presented briefly here.

Skinner defines two general classes of behavior - respondent and operant. There is considerable disagreement as to whether or not these forms of behavior actually differ or whether they are on a continuum, but both forms of behavior may be conditioned by reinforcement. In respondent conditioning a neutral stimulus is presented simultaneously with an unconditioned stimulus which is one that has a high probability of eliciting a response. When this

response occurs in the presence of a reinforcing stimulus, the neutral stimulus also becomes attached to the response by response generalization. The neutral stimulus is now said to be conditioned, but it is subject to extinction by elicitation and must be reinforced further by presentation in close temporal contiguity with the unconditioned stimulus which elicits the desired response. This is known as respondent conditioning and has its historial basis in the work of Pavlov.

The process used in operant conditioning differs considerably from that used in respondent conditioning. The organism's behavior is observed and when the desired form of behavior occurs it is reinforced immediately. This is the form of conditioning used when animals are taught to respond in a manner pre-determined by the experimenter. For example, a hungry rat may move about randomly in a cage. Eventually it comes in contact with a lever which releases a good pellet. The rat eats the pellet and an increase in bar-pressing behavior is observed.

It is not always possible to categorize behavior as being either purely respondent or purely operant. Green (1962) observes, however, that respondent behavior can usually be described as involving Pavlovian or classical conditioning and is closely related to emotional behavior, whereas operant behavior is more important when learning

involves instrumental conditioning, trial and error learning, verbal conditioning, motor learning, problem solving and concept formation. This may be, to some extent, an oversimplification but it appears to be useful in a study of programmed instruction which Skinner claims is a matter of operant conditioning.

Strengthening any kind of behavior depends upon reinforcement. Such reinforcement may be described as primary, secondary or generalized. Primary reinforcement involves a state of affairs that reinforces without need of prior training. Secondary reinforcement derives its effectiveness from a previous process or learning event. Generalized reinforcement comes about through the association of stimuli in the presence of reinforcement. That is a neutral stimulus is associated with several primary reinforcers under a wide range of deprivation conditions at different times. Thus a broadly based secondary reinforcer becomes a generalized reinforcer.

A further distinction which must be made regarding reinforcement is to state whether they are negative or positive. Behavior can be strengthened by the presentation of a stimulus, as when a hungry rat is given food; or by the removal of a stimulus as when the rat terminates shock by pressing a lever. Negative reinforcement is therefore a matter of "escape" or "avoidance" conditioning. Avoidance involves the use of primary negative reinforcers and

escape involves the use of secondary negative reinforcers. The apparent distinction between negative and positive reinforcement seems to be that the former is aversive to the organism while the latter is pleasurable. Unfortunate hedonistic tones can be heard in the latter part of this statement, however, and these must be ignored.

Skinner (1954, 1957) has based his theories of programming instruction on the basis of operant conditioning. He states that machines - especially Skinner-type machines - operate on the principle of developing behavior by reinforcement of successive approximations. This is not the same as rote learning by means of a memory drum because the pupil never gets the same stimulus twice, but in programmed instruction he is actually forced to think out the answer. By the use of reinforcement the experimenter can shape up a desired pattern of behavior and by making use of appropriate schedules of reinforcement he can keep this desired behavior at strength, the only limitation being possible physical exhaustion on the part of the subject.

> "Many of these effects would be traditionally assigned to the field of motivation, although the principal operation is simply the arrangement of contingencies of reinforcement." (Skinner, 1954, p.87)

These contingencies of behavior may be useful not only in simple factual learning but may even be used to produce complex social behavior. Such a program, however, must be carefully designed to include changing contingencies to meet the needs of the learner. Thus a particular type of

behavior may be produced by presenting a sequence of events or stimuli which will produce an action, or a very complex sequences of schedules may be constructed which will produce extended behavior patterns. For example a pigeon may be presented with stimulus pattern A and rewarded by presenting stimulus pattern B. This may lead to a performance producing stimulus patterns appropriate to schedule C and so on.

Skinner suggests that, though experimental evidence has been adduced by means of animal experiments, many of the experimental results can be applied to improving classroom instruction. He concludes that teaching machines and programmed instruction are able to obviate some of the shortcomings evident in our modern education system -- namely, the use of aversive stimulation to control learning, infrequency of reinforcement and lack of skillful programming. As a result of these shortcomings in current educational practice, subjects are not learned quickly or well. Programmed instruction, however, makes reinforcement contingent on the desired behavior and controls the learning process step by step acting at once as guide and motivator to further action. Explicitly, how does this process work?

First sources of reinforcement must be tapped. The major sources of reinforcement in a school situation are found in the subject matter itself. Failing this, other reinforcers must be developed, as, for example, the use of competition or generalized reinforcers often referred to

as extrinsic motivators.

Secondly, these reinforcers must be made contingent upon the desired behavior. Complex behavior consists of very small steps, each of which must be reinforced when it occurs in the correct place in the sequence and each of which must be maintained at strength once established. Skinner points out that since we know little of how schedules of reinforcement operate in human behavior, it is probably safest to provide for small, easy steps and for continuous reinforcement. Since the human experimenter through personal mediation cannot effectively create the best schedule, a machine or device is necessary to perform this function. "The simple fact is that, as a mere reinforcing mechanism, the teacher is out of date." Providing adequate reinforcement is, however, simply accomplished by use of a machine or electric device. Such a device also makes it possible to present a very logical sequence of events, the learning of one dependent on having learned a preceding one. Should such reinforcement prove to be inadequate, supplementary sources can be added in the form of external rewards or competition, which will not detract from reinforcement already present in the program.

The problem which arisesout of this argument and about which Skinner must be questioned seriously is that of scope. Can programmed instruction really produce a variety of behavior such as we desire in our pupils, or will such instruction lead to a reduction in creativity

and divergent thinking?

When creativity is discussed the problem of transfer naturally arises. Kendler (1959) points out that in this regard mediation must be taken into account. A study of transfer is needed to determine how simple verbal responses can develop into abstract ones and how different responses are re-combined to produce novel behavior. Novelty and mediational processes may very well be related, and word-association originality may facilitate problem solution. The problem with programmed instruction may lie in the fact that we reduce variability and thereby increase con-creteness. This may reduce the possibility of association and thus the likelihood of creative thinking. Nevertheless, once these factors have been submitted to research, it appears likely that programs could be written to provide for them.

A second problem arises out of the preceding dis-cussion: that is, the Skinnerian concept of the learning paradigm. Skinner has based his thinking on the paradigm of free operant conditioning and this appears to be too narrow a concept. If, as Zeaman (1959) suggests, this is too narrow a paradigm to include the wide variety of learn-ing activities engaged in by humans, it follows also that the Skinnerian type program will not be adequate. It is impossible to present a paradigm which will be sufficiently inclusive, but any limitations which this may place upon

programmed instruction are only temporary. Since programmed instruction has been successful in those areas in which it has been used, it is likely that the only limitations which need to be considered are the limitations of the programmer.

Procedural Variables The teaching machine program is the basic factor in instruction. Essentially, modern investigators have subscribed to the Socratic method in construction of programs and have tended to follow the Cartesian method of ordering learning events. The method is therefore one of arranging questions and eliciting answers in some prescribed order. The content of the learning task is analyzed first and then the pupil proceeds from simple to complex items until he has grasped an understanding of the concept which he is studying. Unfortunately, the modern programmer has little research evidence available which makes it impossible for him to follow some logical set of rules in ordering the frames of the program so that he will ensure maximum progress as the student moves from the simple to the complex. Principles of programming are now in the embryo stage and an interim approach to programming is necessary in order to carry forward initial research. Such an initial set of rules is probably best generated from an examination of present learning theory. That is, it should be possible to produce some guide to programming which will make use of research evidence collected in other

areas of psychological research. A brief analysis of this research and of its implications for programmed instruction follows.

Types of Programs - At present there are two basic types of Programs: the linear type and the intrinsic or "branching" program. The linear program is the type developed by Skinner and is composed of small steps leading logically through a body of subject matter. The increments of information from frame to frame are small and it is expected that as a result the error rate will be small. Learning takes place when behavior is conditioned to the elements within a frame. The amount of redundancy in the program determines the error rate and the degree of reinforcement is dependent upon the error rate. Green (1962) has given an operational definition of reinforcement which states that reinforcement occurs when the student learns by answering an item correctly. With this definition there can be no quarrel.

The second basic type of program, the intrinsic program, is based on the belief that the steps need not necessarily be small and that consequently the student need not always be correct. The student is presented with a problem and with a group of multiple choice solutions from which he must choose. When the student chooses an answer he is instructed to move to a specified frame. This frame then tells him whether his answer was correct or incorrect

and, if incorrect, it explains why. The frame then may instruct the student to return to the original item which he had answered incorrectly for another trial, or it may direct him through a branch, for example, to improve his knowledge of a specific concept. He is then allowed to proceed via the frame which he originally missed or to a frame which continues the series. The intrinsic program differs from the linear program in two major respects: First, the "intrinsic" program uses a different feedback procedure and secondly, it makes use of recognition rather than recall. The "branching" program has some resemblance to a multiple choice test, but the immediate feedback device prevents the student from learning the wrong member of a set of answers. Crowder has objected to the linear program on the grounds that the small step is an insult to the intelligence of the learner. It may also be true that there is little reinforcement from learning small steps which seem insignificant to the learner. Perhaps there is a need to be wrong sometimes and perhaps there is some form of generalized reinforcement which operates here so that there is incentive arising out of accomplishing something difficult. This is a matter which must take individual differences in ability and experience into account. Crowder (1960) summarizes intrinsic programming as: "The technique of using a student's choice of an answer to a multiple choice question to determine the next material to which he will be

exposed." A number of adaptations of the method are possible. These are summarized below. The major feature of this type of programming is that it makes possible feedback control, which Crowder claims is superior to the technique of supplying knowledge of results.

1. The simplest type of intrinsic program is the one in which each wrong answer refers the student back to the original choice page to try again.

2. A second type utilizes a simple subsequence. In this case the program digresses but returns the pupil to the frame at which he made the original error.

3. If the pupil is taken on several detours for review and is then returned to the main stream, a wash-back program is being used.

4. A wash-ahead program may also be used. In this case a group of easier subsequences leads slower learners to the next major step.

5. The most complex variation of intrinsic programming is referred to by Crowder as a complex wash-back program. Here the errors are weighed and the pupil is returned to some previous frame, usually from 0 to 3 steps back depending on the extent of the error made.

Once the programmer has decided which basic type of program he intends to use, he must also decide on a procedure to follow in programming. It is here that the pupil who is to use the program must be consulted. It is the student who must work through; and, in the absence of evidence on this point, it is the problem of the programmer to write a program that can be tested pragmatically. A number of approaches to programming have been suggested. Only one such approach will be reviewed here and that is one known as the Ruleg System developed by Evans, Homme

and Glaser (1960). They suggest the construction of a "ru" matrix on a horizontal and vertical axis. This matrix consists of all the concepts or rules necessary in the program. Using this matrix as a guide the programmer collects examples of concepts in all their possible interrelationships with one another. These are ordered in logical sequence and then the programmer proceeds to write examples that will define these concepts. The likelihood of transfer and the extent of generalization will be determined by the number and appropriateness of the examples. Discrimination may be aided by the inclusion of negative examples of the category being examined.

Supplementary Devices - Supplementary devices may be used to make the program clear and to give it concreteness. Such devices may be in the form of a panel which usually consists of a graph or table or some information to which the student may refer. Such a device may be considered a prompt because it tends to make the desired response more likely. In addition, the programmer may use formal prompts which are usually direct stimuli included to elicit a specific response and are in the same form as the response - for example, a textual prompt whereby the student sees a word in the presented material which elicits the correct response. If it is desired that the student be responsible for the response to a greater extent, a thematic prompt is used which is less direct than the formal prompt and is

usually considered to be a hint. Other types of prompts include sequential prompts, the use of models and emphasizing of important details. Another supplementary guidance technique has been developed by Holland (1960). This technique is known as "fading" or "vanishing". It consists of presenting a picture, chart, diagram or series of items which the pupil is expected to learn. The stimulus is presented a number of times, but each time some of the information is missing and the student must supply it. Eventually he must supply all of the stimulus. This type of prompting is used especially in the teaching of spelling and foreign languages.

Task Factors - If a body of subject matter is to be learned effectively, it is necessary to analyze that particular body of material. To do this the steps involved in learning the material must be described or classified in some logical manner. Cotterman (1959) is working on a model for such a classification. His classification is made using interaction between task characteristic and practice variable as a differentiating criterion. By this method of classification it is possible to determine whether or not a specific mode of instruction can produce significant differences in learning particular types of tasks. After this has been done the various types of tasks can be examined to determine critical differences among them, and new tasks can be classified accordingly and subjected to

treatment best suited to learning that particular type of task.

Stolurow (1961) has suggested one set of features which could prove useful in determining the method by which learning certain tasks might be handled by the teaching machine program. These will be reviewed briefly here.

Number and sequence - It is not yet clear which sequence of programming is the most effective. Cook and Kendler (1956) have suggested that the prompting sequence, in which the learner is given a cue stimulus and a prompt before he makes the response, is superior to the confirmation sequence, in which the learner is given the cue stimulus and then is expected to make the response which is then confirmed. The relative effects of these methods have not yet been thoroughly investigated. A further problem involves the use of the stable sequence (serial learning task) as opposed to an unstable sequence (paired associates task). Again there are advantages and disadvantages to habit systems developed by means of either procedure. The stable sequence is learned more quickly but it is less adaptable. Since it is a habit learned in a certain order, any changes in that order need to be learned as separate tasks. It follows, then, that the sequence used will depend upon the material to be learned as well as upon the type of learning that is desired by the programmer. If broad principles are to be learned, the programmer will proceed to vary the

items in various ways. If, however, memory type material is to be mastered, redundant items in the program will consist of identical stimuli and response items.

S and R Limits - If one follows the Guthrian learning paradigm, it becomes important that the learner be presented with identical stimuli if the proper response is expected to recur. It is well known, however, that identical stimuli are almost impossible to produce; and it follows that learning will be impeded to the extent that stimuli which are to produce identical responses differ. At the same time, learning will be accelerated if the response to certain stimuli is permitted to vary over a rather broader range. It also follows that stimuli that are alike should not be grouped too closely together as confusion may result because of the difficulty in discrimination.

Meaning - In order that two independent experiences be linked together, it is important that mediation take place. The stimulus must be arranged carefully if this is to happen. To make sure that useful responses are available in a learner's repertoire, it seems that in programming it would be wise to build up a "web of association". The transferability of associations is determined by two major factors: membership in a class and orderability. By membership in a class we mean that a stimulus can be classified on the basis of previous experience and used in a new one. By orderability we mean that there is some inter-

dependent relationship among a series of items and that, as in counting, there is established both remote and adjacent association. Retention is thus aided.

S-R Linkage Pattern - The learning task may be used to create a one-to-one relationship, a one-to-many relationship, a many-to-one relationship or a many-to-many relationship. If each step in the learning process is analyzed on this basis, it follows that a much more definite learning task will be accomplished.

S-R Homogeneity and Variability - If the stimulus and response sets come from the same population - e.g. words are paired with words - the S-R linkages are said to be homogeneous. If they come from different populations - e.g. pictures are paired with words - they are said to be heterogeneous. Some research evidence (Lumsdaine, 1949) suggests that heterogeneous sets are learned more quickly. Stolurow (1961) suggests that this is because interference may take place in a homogeneous set where discrimination is not complete. The same condition may occur when the sets of stimuli and responses are incongruent or incompatible, and the differences from item to item in a set do not vary in the same way.

A further difference not discussed by Stolurow is the relationship between task length and meaningfulness. It is rather obvious that the length of the task will effect the meaningfulness of the material and that the ability to

deal with a given concept either by whole or part method will depend on the ability of the individual pupil.

Practice Conditions - Adaptive machines are often evaluated on the basis of their ability to wait for the learner's response. Briggs, Plashinski and Jones (1959) compared the effects of self-pacing and machine-pacing. They used a paired-associates task and found that there was little difference in performance. Since paired-associates learning is less affected by distributed practice, this experiment should be repeated with a serial-learning task. Further evidence on this subject is not available.

Much work needs to be done in determining the various effects of certain other practice conditions. For example, we need to know more of the relative effects of stimulization by means of various modalities; we need information on the order which is best in presenting a stimulus and response. Do we move from the most familiar to the least familiar? In what kind of material does it work better to have the simpler part of the link in the response position? It is also necessary that we analyze a response in terms of the stimulus which can produce it. The stimulus must meet at least three criteria. It must be reliable - i.e., it must produce the desired response. It must elicit the response at optimal speed. If the response latency is kept at a regular rate, a response set may be formed which will

pace the student instead of having him do the pacing. A third criterion of a response may be its transferability. For example it is possible that with the Skinner-type program the responses are so carefully arranged for that transfer may be impeded. The Crowder type of program presents the material in rather larger blocks, making it necessary for the student to do much of his own "associating" or transferring. Because he has done this, the student is now able to make the necessary conversion to motor action or to new material more easily.

Feedback Factors - Feedback usually takes the form of stimuli which provide reinforcement to the learner. Reinforcement is any event which follows a response and which makes that response more likely upon the recurrence of the original stimulus. The effects of reinforcement may be considered more broadly if the feedback has the effect of re-directing behavior in addition to reinforcing it by telling the learner that he is correct. This is usually considered to be a prompt. Skinner has pointed out that the regular instructor in a classroom cannot possibly produce continuous reinforcement and argues that teaching machines make it possible to provide such reinforcement. There is, however, little proof that this is the best type of reinforcement schedule; and it may be that some type of mixed strategy of reinforcement is more beneficial, particularly as it relates to retention.

The importance of immediate feedback cannot be denied. The effect of immediate knowledge of results on learning has been studied by a number of people. Pressey (1950) gave a group of college students a series of tests using a punchboard. This made it possible for the students to know immediately if they were right or wrong. Another control group took the same tests in the conventional way and then had them scored and returned the next day at which time they were allowed to discuss the errors they had made. On a criterion test it was found that these two groups did equally well, even though the punchboard group were not given a discussion period. Pressey concluded that knowledge of results makes learning more effective. This was true even when items were so constructed as to demand transfer from practice materials to new items.

Time Factors - Closely related to feedback factors are time factors. Green (1962) suggests that there are two major types of conditioning that relate to time. The first of these we know as "cramming", which is achieved by continuous reinforcement of tightly massed material. Such a schedule will result in an increase in response probability but less retention. A second type of conditioning we know as maintenance of behavior which may best be obtained by scheduling reinforcement or by spacing training sessions. The result in this case in reduced response ratio and better retention. Since there may be interaction between

retention and the type of task performed, it is probably more appropriate to combine these two approaches when programming instruction.

Stolurow (1961) also discusses critical time intervals. Efficient learning requires minimum intervals between the cue and the response and between the response and the reinforcement. Optimum time intervals are required between the reinforcement and the next cue stimulus and between practice periods. The latter time interval will be a function both of the material to be learned and of the learner. A further confounding issue is the problem of learning set which seems to cut across the effects of time interval. This could easily be handled in programmed instruction by preceding a unit with a section which might help to bring about this set.

When should knowledge of results be presented? Stephens (1953) used the Drum Tutor and the punchboard to try to determine the best time to provide feedback. One group used the Drum Tutor to take a test after learning the material by conventional instruction while the other group learned the material by practice with the punchboard. Those who used the Drum Tutor showed superior learning indicating that early guidance, that is guidance provided within the frame so that the possibility of error is reduced, may be an important variable in the learning process. Should this finding be borne out it would provide

an argument for more elaborate teaching devices and extensive programming to make possible the reinforcement of each step of original learning. This would support the position of those who use Skinner-type programs. On the other hand, if immediate knowledge of results effected by a testing device some time after learning had taken place, can be demonstrated to be equally effective, then a more economical means of increasing the effectiveness of instruction can be invoked. Extensive programming would not be required. The latter would also tend to support Pressey's opinion that such devices should be used as aids to teaching and should not be used to teach all the material in a course. Again, it is the contention of this paper that the best time for feedback will be a function of the type of program and the practice mode that is used.

Another phase of the feedback conditions may involve the problem of guidance and prompting. Briggs (1962) reports that in a number of studies it has been found that the error rate is directly related to the amount and conditions of the guidance given. Guidance may take three forms: cue guidance - the isolation of critical clues; response guidance - the deliberate elicitation of the correct response; and relational guidance - the presentation of the principle to be used to achieve correct performance. The programmed material contains implicitly the factors needed to produce the response and the re-

inforcement schedule must be so arranged that the proper response is selected in accord with the criterion. A number of ways of giving guidance have been developed. These do not enter into the problem of this thesis and are therefore not discussed here. It is necessary, however, to look carefully at one final point in this discussion and that is error rate.

Skinner has argued consistently for errorless practice in programmed instruction. For this reason programs that follow his model use small easy steps with continuous reinforcement. The question arises as to whether or not errorless practice is necessary or desirable. In the first place, it would appear that the program should be evaluated not on error rate alone, but on other factors such as retention and transfer. In the second place, it is necessary to take a much closer look at the motivational aspects of the case. It may be that the bright student or even the typical student is not challenged by the small step program and that he does not really discriminate thoroughly unless at times he is wrong and is then shown where he is wrong so that he may adjust his understanding accordingly. This is similar to Bruner's (1960) contention that we give both positive and negative examples in order that discriminative learning may take place in combination with generalized learning.

Response Conditions - The method by which the pupil is

expected to respond to the program is important in the learning process. An analysis of this problem entails the study of three interacting approaches to programming; the use of correction versus non-correction procedure, the type of response mode employed and the practice mode used in programming.

Pressey and Skinner have both employed devices which require the learner to repeat those items in a program where he made an error. Briggs (1949) compared the use of the punchboard with and without correction procedures. He reported that the correction procedure was superior in terms of test results and that both methods were superior to regular testing in which there was no provision for knowledge of results. The procedure involved here led investigators to develop various modes of operation herein known as presentation modes. The first such investigation was made by Bryan and Rigney (1956). They used the Tear-a-Tab device in which the learner pulled off a tab which he thought contained the right answer to a multiple choice question. Under the tab appeared the words "right" or "wrong". If the wrong choice was made the learner continued to pull tabs until he made the right choice. A control group answered the questions in the usual way and a third group also pulled tabs like the first group but in addition to finding the terms "right" or "wrong" under the tab, they also found an explanation why they were wrong.

A multiple choice test revealed that the group getting the explanations was superior to the other two. From this point on a number of practice modes have developed. These modes are defined and discussed in Chapter III.

Irion and Briggs (1957) reported that the most efficient learning took place when the quiz mode was used. In this method of operation the learner reads the item and then asks the machine for the correct answer. This method proved to be more effective than the practice mode, in which the learner responds until he gets the correct answer. It was also more effective than the single-try mode, in which he is informed only that he is right or wrong and then goes on to the next frame. The study of Irion and Briggs suggests that the interval between cue stimulus and response may be critical andin the quiz mode the prompting conditions present - i.e., surrounding material - may be vital to learning. If the findings of Irion and Briggs can be substantiated, relatively simple teaching devices may be effective.

The final response condition in which we are interested is the practice mode. The learner is expected to make either an overt or covert response to a stimulus. Evans, et al (1959) used two groups of undergraduates to give a course in "Fundamentals of Music". One group went through the material constructing one or more responses at each step, while a second group were instructed to make no overt response. The latter group took less time to complete the

sequence and also achieved a better score on the criterion test. Porter (1957) claims that if no overt response is required, the device used cannot be called a teaching machine at all. This reasoning is also followed by Skinner, who asserts that the pupil must make an overt response in order that the learning will be fixed. Crowder and others disagree and say that the constructed response is only useful as a measure of what the student has learned. Siverman and Alter (1961) found no difference in learning when they compared a "response" with a "no-response" group, both of which had taken a 90 item program dealing with the binary number system. Kaess and Zeaman (1960) investigated the effect of overt wrong responses. They found that when a wrong response was overtly made, it was strengthened, but when a wrong answer was merely read as a part of the instructional program, it was not strengthened and could be easily extinguished by subsequent non-reinforced elicitations. This led to the conclusion that the amount of associative strength formed depends on the practice conditions. In the production of correct responses, however, this did not hold true and reinforcement of overt responses did not appear to be any more effective than reinforcement of covert responses.

Other investigators, (e.g. Holland, 1960) have found that where overt responses were not required, error rate rose, presumably because construction of the correct answer

in a constructed-response program drew attention to critical information contained in the frame. No clear evidence is available on the subject, but evidence collected by Evans, et al. (1959) would indicate that, in any case, a blank does not need to appear in every frame of a program if progression from frame to frame is clear and systematic and if feedback is adequate.

Other investigators, (Gotkin, et al. 1962; Roe, et al. 1960; and Silverman and Alter, 1961) fail to find any difference between response modes used. Basically the controversy over covert versus overt response is a theoretical issue in learning theory with its genesis in the S-S and S-R controversy. The S-S theorists attach little importance to overt responses, preferring to emphasize perceptual organization and central cognitive processes. The S-R theorists, on the other hand, feel that a position which emphasizes the connection between stimulus and knowledge leaves a gap between stimulus and action. More recent S-R theorists such as Hull, Mowrer and Osgood, have attempted to close this gap between stimulus and action by postulating a central mediating process in which the overt response is a major factor.

Since research findings regarding the efficacy of covert responses disagree, it may be concluded that there is some interaction effect which clouds the results of experimental study. Walker (1962) suggests that such interaction may exist between the type of material used and

the conditions of overt response. Such a postulate could be extended to include other variables such as method of presentation and intelligence level of the pupil. In fact, the issue may not lie in the problem of covert versus overt response at all, but may resolve itself by a control of temporal contiguity patterns. That is, differences may be due to the fact that it simply takes longer to make an overt response.

> "Thus, with covert responses the interval between the stimulus-response connection and the interval between the cue-stimulus and the next are shorter. On the other hand, the effectiveness of overt practice may lie in the degree to which the overt response is cued by the otherwise practiced covert response. If the student who can think a word can write it correctly, there already exists an established cue-response relationship between thought and action and it would presumably make little difference whether he was required to make a covert or an overt response in the course of learning. On the other hand, if the student had not learned this relationship then it might make a great deal of difference if he practiced an overt or a covert response, for in practicing the overt one he would be building up the required relationship between saying the word to himself and writing it." (Walker, 1962, p.423)

If, then, the student is to make use of responses that are already practiced and a part of his repertoire, covert practice will be sufficient.

Methods of Presentation - Roe (1960) compared the effects of teaching machines, programmed texts, programmed lectures and a standard lecturer. The program used a 192 frame sequence on elementary probability. He preceded the actual study with a pilot study to check the comprehensibility

reliability and validity of the programmed materials. The subjects were 186 freshmen engineering students at the University of California. They had previously taken an engineering aptitude test which was used to divide them into four ability levels. The modes of presentation were: (a) two types of teaching machines -- a Skinner-type constructed response and a multiple-choice response; (b) two types of programmed textbooks -- one requiring response and supplying reinforcement in the form of the correct answer on the next page; the other presenting the items as in a conventional text and requiring no overt response; (c) two types of lecturers -- lecturers who followed the sequence of the program and a conventional lecturer. The presentation of the program was followed by a test consisting of 20 free-response items. Some interesting results were obtained. All groups receiving programmed instruction performed better than those receiving conventional instruction. Unlike many other studies, however, this study reports that students in the lecture groups took less time to complete the course than did those using programmed texts or machines. Those who did not need to construct answers also took less time than did those who were required to construct or choose responses. The students in the lower aptitude groups did not score as high as the students in the upper quarters, but there was no indication that any one of the methods was particularly better for a particular aptitude group on the

basis of either the criterion variable or learning time.

This report is well written and the study well designed. Perhaps the major problem in such studies is the length of the programs. It appears likely that many important variables are not permitted to function when the program is too short and replication with longer programs seems necessary before too many conclusions are drawn.

A final study comparing programmed and conventional instruction will now be mentioned. Ingles (1961) used 112 men taking a company course in customer engineering to give the first 15 hours of an IBM 7070 computer training course. The program was presented in the conventional classroom method (lecture-discussion) and by means of a 719-frame programmed textbook. Two classes of 21 students each were designated "control classes" and given the company's usual lecture and discussion course. In each of the three subsequent months other classes undertook to learn the same material from the programmed textbook. At the end of the course the students took a comprehensive two-hour examination and filled out an attitude questionnaire.

The author of the study reports that the students who worked with programmed texts required 11 hours to complete their work. He does not state if this was average time but we note that it took the regular classes 15 hours.

Students who worked with programs achieved a mean score of 95.1 on the criterion test while the conventional classes achieved a mean score of 86.2, a difference significant at the .01 level. The attitude questionnaire filled out anonymously was favorable toward programmed instruction. Although this study appeared to be significant, it will be noted that the experimental groups were not run at the same time as the controls. The author also admits that only educational level among the groups was controlled statistically. Furthermore, the group was a select one and replication is needed to obtain evidence about average groups and perhaps children in particular.

Criticism of Contemporary Studies in Programmed Instruction

Although research on programmed instruction is for the most part still in the preliminary phase, some criticisms regarding design and analysis of data are already apparent. Most of the studies to this point have been exploratory. Controls are often left out or ignored and study has been limited to individual variables relating to the program, the learner or the criterion data. Robinson (1961) has made some insightful observations regarding the research literature. He points out firstly that the effectiveness of programmed instruction is still in doubt. Though experiments, such as the Roanoke experiments (1959), in which the experimental group learned mathematical concepts in half the time required by those taking conventional instruction, have shown the effectiveness of teaching

devices, it is not yet known what the long-term motivational effects are. Control in these situations is at best theoretically effective because it is well known that the pupil and the teacher in any classroom situation can certainly not be standardized.

He points out, secondly, that there is little research regarding the role of the teacher in programmed instruction. Most researchers have been led to accept without question Skinner's belief that the teacher cannot provide small steps, a satisfactory sequence, feedback. This, of course, may not be true and it may be that what teaching machines and programmed instruction will give us will be a well-planned instructional program to be used by the teacher in a conventional presentational manner.

Thirdly, he discusses the characteristics of an optimum program stating that "--there is no proof that 'overt' response is superior to 'covert' response, that the error rate of the learner should be extremely low, that there should be feedback on every trial, and that it should always be positive feedback--". No program type has yet been proven superior for all subject areas and comparisons of a variety of programs in a specific subject field are not yet available.

Finally, he challenges the adequacy of existing theory which he feels has, in this area at least, been reduced to a theory of operant conditioning not suitable

as a description of human learning in all its phases. It is fairly easy to see that Skinner's approach may be inadequate in many complex areas such as geometry for example, in which there appears to be little justification for speaking about "rate of response" or "successive approximations".

Gilbert (1960) goes much farther than any of the other critics of research and states that an about-face will be necessary if progress is to be made. He holds that present research merely shows that there are individual differences in pupils and programmers. Improvement, he suggests, must come in two areas: First, improvement of a given educational subject matter by having an expert teacher practice writing frames for a pupil and ignoring the machine until the program has been tried. (This probably means that the real question is one of how to teach a specific subject.) Secondly, if human learning principles are to be discovered, expensive laboratories will be needed in which subjects can be used who will act as their own controls over long periods of time. Only in this way will systematic knowledge be obtained, and the danger of following up variables of pre-conceived importance be avoided. While Gilbert's theory is still somewhat heavily dependent on the theory of operant conditioning, it does appear to offer important suggestions for the improvement of research.

CHAPTER III

THEORETICAL FRAMEWORK

Definitions

Automated Instruction - This term is broadly used to include any method whereby the teacher function is provided by a wholly or partially automated sequence of instructional segments prepared in advance in the form of a program. This term is also used to include the narrower concept of programmed learning which refers more specifically to the frames of a program than to the method of presentation as a whole.

Construction - The process of requiring the subject to write orprepare an answer rather than requiring him to choose one of several alternatives.

Confirming Mechanism - A device by means of which the subject is informed of the correctness or incorrectness of his response. This may be by means of some auditory or visual stimulus, or, in the case of constructed answers, by requiring the subject to compare his answer with a model supplied by the programmer.

Error Rate - This term is used to refer either to the number of errors made by a specific student on a total program, or to the percentage of a given group giving wrong answers to a specific item in the program. This is the criterion usually used to determine revisions in the program.

<u>Fading</u> - The gradual withdrawal of stimulus support in presenting items. (See also "vanishing.")

<u>Feedback</u> - The process by which information is given the subject which makes possible modification or correction in the pursuit of a sequence. Since it is a corrective device, it is considered as a variable distinct from reinforcement. It is an indication of the extent of the correctness or incorrectness of a response and may vary in the amount of information it gives and in the temporal delay involved between response and feedback.

<u>Frame</u> - Ordinarily a single item of information is exposed at a time. Such an item of information is designated as a frame in the Skinner-type program. Recent developments in programming, however, have made necessary re-definition of this term and it is now used to include any material presented to the subject at one time. This makes for greater flexibility in programming and makes possible a greater variety of means for handling individual differences both in subject matter and human abilities.

<u>Hints</u> - Devices used within the material presented to make the correct response more likely. At present this term is used interchangeably with the term "prompt" which might better be reserved for the type of guidance which consists of added material.

<u>Operant Behavior</u> - This is a fundamental concept in Skinnerian learning theory and refers to responses which

operate on the environment to bring about desired changes.

Presentation Modes - In this paper presentation mode will be given a broad interpretation to indicate the method of practice employed by the subject to learn the original material. This may be thought of as being distinct from an instructional method which term might be applied more accurately to a combination of method of practice and programming itself. Irion and Briggs (1957) have identified four modes:

1. Quiz mode, in which the subject reads the item (frame) and then "asks" the machine for the correct answer. It is presumed that a covert response is made to the stimulus item.

2. Modified quiz mode - The learner chooses an option, being presented at once with the next question if correct and, if incorrect, an error signal is accompanied with a direct prompt that shows which option he must then choose to be correct. In the constructed response type of program, the pupil constructs an answer, "asks" the machine for the correct one, and then goes on. In a modification of this program, he might be asked to re-write the answer if he made a wrong choice.

3. Practice Mode - The learner is asked to keep choosing options until he chooses the correct one. This type of program is most suited to a multiple choice type of frames.

4. Single-Try Mode - The learner can respond only once to each problem. Some method is used to indicate whether the right or wrong choice has been made, but the subject is not permitted to make another choice, nor is he given any further information as to the correct answer.

In addition to the presentation modes suggested by Irion and Briggs, other forms are available which may be useful for purposes of experimental comparison.

1. Test Mode - The learner writes answers to questions which deal with the material which has been presented previously in some form. The subject has no knowledge of whether he is right or wrong and reinforcement does not occur until the papers have been scored by the examiner. It is to be noted that this process is in many ways similar to conventional classroom instruction.

2. No-Practice Mode - In this type of presentation the subject merely reads the information as he would in a textbook.

Panel - A short passage of prose material, graphs and similar material presented or studied along with a program. This was originally used together with the Skinner Disk machine.

Program - The program consists of the subject matter to be learned by the subject through means of the teaching machine or some other device. There are two basic forms of program:

Intrinsic (Branching) - in which the program is so constructed that the erring subject is directed along corrective pathways before he is permitted to proceed to the next step in the program. This type of program is possible only where multiple choice answers are available.

Linear - in which the pupil must construct a response. The items of the program are presented in small-step form and require that all subjects follow through the same pattern of stimuli and responses regardless of ability.

Response Mode - This refers to the form of response made by the subject in working his way through the program. Two possibilities exist:

Overt vs. Covert Response - In the former the subject must make some physical response to indicate the answer he has chosen. In the latter no specific action is required, but it is assumed that the pupil is aware of the correct answer and may be making implicit responses in one form or another. In other words, he is "thinking".

Constructed vs. Multiple Choice - The subject may be asked to construct an answer or merely to choose one from a group of alternatives.

Scrambled Book - A device which permits the student to record his response in programmed sequential form, but in which the student is directed to different pages not necessarily in consecutive order. By means of alternate choice responses at each step, branching to new or review material is made possible.

Step Size - This is a measure of the increment in difficulty level between the steps of a program. Large steps may provide greater challenge and make the program shorter. Small steps require a longer program, but reduce the error rate, thus making for continuous reinforcement.

Vanishing - The program itself makes possible the elicitation of an appropriate response in the presence of certain stimulus. In order to strengthen the desired behavior the cue which gives rise to a desired response may be repeated several times to insure thorough learning. It may then be gradually faded until the desired response is elicited in the presence of minimal cue material. This is an aid to both discrimination and generalization.

Theoretical Framework

It has been mentioned that Skinner has consistently argued for a small-step program in which the pupil constructs his response. The contention is that the constructing of the response fixes the answer in the mind of the pupil. It follows that the programmer must arrange the frames so that wrong responses are avoided and therefore not fixed in the mind of the learner. The procedure advocated by Skinner has some serious limitations. In the first place, the long-term motivational effects of small steps need to be investigated. Though reinforcement is continuous, there is little to reinforce if the pupil is left unchallenged by the small increment of learning that he feels has taken

place. Secondly, because of the limitation imposed by the machine-scoring of items -- even when the student makes the comparison to the correct answer -- the scope of the response learned may be very narrow and the strength of the associative bond between stimulus and response limited. Thirdly, there is no proof that low error rate is always desirable because discrimination can only be effectively achieved when closely related examples and negative examples are presented to the learner. In other words, synthesis and analysis depend on a critical examination of material, and this may not be possible without having the learner make some errors. Finally, one crucial test of whether or not material has been learned is long-term retention. There may be some connection between error rate and long-term retention. That is to say, it may be that pupils will retain material more effectively if they made some error in learning it which drew the crucial material to their attention or which made them refer to the principle underlying the material that was being learned. This area, however, remains to be investigated.

Crowder has opposed the theories and linear program recommended by Skinner and has devoted his efforts to the development of the intrinsic program. This program makes it possible for each pupil to gain additional information in an area in which he is weak by following a sub-program or "branch". Some allowance is thus made for individual

differences. Two serious problems limit the usefulness of this technique. First, the program is built on the pattern of the multiple-choice test, and this means that there is a danger of making recognition the criterion of achievement. It is feared by some that understanding will be limited if recognition is the only criterion towards which pupils are asked to work. Second, the possible number of responses is limited to those the programmer has decided to include in the options of each frame. That is, the programmer defines the branches of the program ahead of time but he has no way of knowing the repertoire of the pupil. This may reduce the possibilities of association and mediation and in fact make this type of program more rigid than the linear type program in which the pupil is free to construct any response.

Once the programmer has decided which type of basic program is suitable for the subject matter he wishes to present, he must make some further decisions. It is important that he know which type of presentation mode he wishes to use. If he is involved in research it is possible for him to choose among at least six of these methods and among a number of variations of each. Within each method of instruction are contained a number of interacting variables which the programmer must consider and weigh carefully before proceeding. Table II indicates the possible presentation modes and also indicates the extent to which some of the important variables related to learning are present in each.

TABLE I

CLASSIFICATION OF INTERACTING FACTORS IN PROGRAMMED INSTRUCTION

	Presentation Mode	Response Mode	Program Type	Reinforcement	Feedback*	Possible Error Response
1.	Quiz	Covert	Constructed or M.C.	Max.	Max.	Min.
2.	Mod. Quiz	Overt & Covert	Constructed or M.C.	High	Max.	Low
3.	Single Try	Overt	Constructed or M.C.	Medium	Medium	Medium
4.	Practice	Overt	Constructed or M.C.	Medium	Medium	Medium
5.	Test Mode	Overt (S-R Temporal)	Constructed or M.C.	Min.	Min.	Medium
6.	No Practice	Covert (S-R Temporal)	Text Book Type	Min.	Min.	Min.

*Feedback

1,2 provide for guidance
3,4 tell if right or wrong
5,6 neither of the above

The quiz mode is placed first on the list because it appears to meet best the requirements of good instruction. There is always the question of the type of response required of the pupil, but it has been noted in the previous chapter that a covert response may be just as good as an overt response unless the pupil is learning a new vocabulary or set of symbols or, perhaps, in the case of young children who as yet have little control over verbal responses. In the case of rote or skill learning overt responses are also probably more effective. The quiz mode is adaptable because it can make use of either the linear program, in which questions are asked in the form of missing parts to be filled in by the pupil, or it can use multiple-choice type items. Since feedback is immediate, reinforcement should be at a maximum. Finally, if it is true that error responses are not fixed if they are not constructed by the pupil, this mode should have definite advantages over those which require a constructed response.

The quiz mode can be modified to require constructed responses. If this method is used it is understood that error responses may be fixed. This factor is vitiated, however, by the fact that feedback in the form of the correct answer is given immediately. The fact that some reinforcement of error responses does take place, however, might make this mode somewhat inferior to the quiz mode.

A third type of presentation mode is the single try

mode. The pupil makes an overt response and is then told whether he is right or wrong. This means that knowledge of results is at a medium level and that reinforcement is reduced considerably, being dependent on error rate. Since the pupil is not told the correct answer, the wrong one may not be extinguished.

The practice mode of presentation requires that the pupil make responses until he is right. The feedback consists of only an indication that the right or wrong choice has been made. Several limitations are apparent. Reinforcement can occur only when the correct choice is made; and since this type of practice lends itself only to some form of multiple choice program, error rate can be high. It has the advantage, however, of providing the pupil with the right answer and depending on the number of choices available in the frame, it may, for this reason, be superior to the single try mode.

A fifth type of presentation is the textbook type of study in which the learner merely reads the material to be learned. He is not asked to respond and for this reason there is limited feedback and no reinforcement except such reinforcement as may accrue from completion of the task involved. It is to be noted that there is also no possibility of error response unless one conceives of this more broadly in the sense that misunderstanding due to lack of discrimination may take place. This would probably be a function of motivation.

The test mode is also similar in some ways to conventional instruction. The student is presented with material which he reads. He is then asked to take a test which is scored by the examiner and returned later, at which time the students may check their work and, if the examiner so desires, a discussion may take place. In any case, though there is feedback, it comes later than it would in the other types of presentation. Since feedback must be given quickly for reinforcement to be effective, this practice mode is less effective and the error rate may be considerably increased. Errors made tend to be reinforced and the effects of such reinforcement are not cancelled out by immediate feedback.

For the most part investigators up to this time have proceeded on the basis that programs can be examined by research involving one variable at a time. It is the contention of this paper that the method of programmed instruction which will result in effective learning is a matter for study by means of examining the interaction effects of a number of important variables related to learning. Not only do feedback factors, subject factors and task factors enter the picture, but the intelligence level of the pupil, the nature of the programmed material and theequipment available to automate instruction are also of importance and research must undertake to examine these variables as they appear in instructional situations -- namely, as a unit each part of

which may contribute to differences in the amount and rate learned.

Postulates

The basic idea of this thesis is that the process of instruction can be considered as a subject for explicit study leading to a technology of instruction and perhaps form the basis for further research in the development of a theory of instruction which will take into account such variables as the structure of a specific subject matter, learning theory and method of presentation. This is not to say that known principles of learning, which have been based to a large extent on animal studies, are not of value. It does suggest, however, that the interacting effects of many of these principles can be determined through the use of programmed instruction and that knowledge thus gained can be useful in improving classroom instruction. On this basis the following postulates are submitted.

I. Retention

Retention is a function of the method of instruction used. Best results will be obtained if reinforcement is at optimum level and if it is given as soon as possible after the response has been elicited. As distinct from reinforcement, feedback is important in guiding the learning process and making reinforcement possible. It directs the learner to make the correct response. Since overt wrong responses may be

reinforced, it is important that feedback be given immediately so that the effects of the wrong response may be reduced. If the feedback is given early in the learning process - that is, during each step of the learning process, error rate will be reduced and the reinforcement ratio will be higher.

II Practice

A. The type of practice which is best during initial learning is determined by the retention criterion:

1. If the retention criterion requires recall, practice with constructed responses should prove to be superior.

2. If the retention criterion requires recognition, practice with multiple-choice type items should be superior.

B. The type of practice response which results in the best learning to a criterion is a function of the type of learning to be done. If the pupil already possesses the repertoire of symbols needed in responding to the presented material, a constructed response may not be necessary. Thus a constructed response should be best suited to such tasks as learning a foreign language and may be superfluous in such subjects as social studies. Age may be a further factor -- that is, a young child might gain more from constructing a response than would an older person for whom responding

in this way is already a motor habit.

III Variability

A. The variability of scores on a criterion test is a function both of the method of instruction and of the type of material under study. If the program consists of steps which are optimal for the subject taught and for the learner involved; if feedback is adequate to keep reinforcement at an optimum and error rate low; that is, if good teaching is being carried out, homogeneity will increase. There is little evidence, however, that this will be true if higher mental processes are involved or if time is taken into account as a factor. It is therefore further postulated that homogeneity will decrease as difficulty of subject matter increases. Brighter pupils, released from the restrictions of the conventional classroom, will produce superior gains when using programmed instruction.

B. The type of program to be used is determined by the aptitude level of the pupils. Pupils with high ability will be able to progress more quickly with larger steps. Less frequent feedback is necessary and reinforcement must be arranged in such a way that motivation is maintained at an optimum level. It appears that a variable ratio of reinforcement is the most satisfactory arrangement for school type tasks. Whether such an arrangement will be adequate depends

on the attitude of the learner both toward the subject being taught and toward the method being used in instruction.

IV Response Mode

The advantages to learning which accrue from the construction of a response are a function of error rate in a program. An overt wrong response is more difficult to extinguish than is a wrong response made covertly. If the pupil sees or "thinks" the wrong answer, but does not construct it, he can more easily re-learn the correct response and extinguish the wrong one -- i.e., the associative strength of an incorrect overt response is greater than that effected by an incorrect covert response. Reinforcement of a correct overt response is, however, no more effective than reinforcement of a correct covert response. (Stolurow, 1961).

V Time

The time required to complete a program successfully will be determined by the amount of activity required of the subject. If the type of material is such that a new repertoire of symbols is not required, a simple teaching device which provides for a small delay between presentation of the stimulus item and the feedback will be sufficient. Any slight gains in learning or retention for subjects who construct responses will

be outweighed by a saving of time for those who respond covertly.

Specific Hypotheses

I Learning is a function of the interaction between the method of presentation, the mode of practice and the mode of response. When the best combination of these variables is used, learning will be at a maximum, other things being equal. The extent of the presence of a specific variable is inherent in the design of the program. This study will compare four of the presentation modes -- namely, the quiz mode, modified quiz mode, no-practice mode and the test mode. The following results are hypothesized.

a) Pupils using the quiz mode will score highest on the criterion test and will cover the work more quickly than pupils using the modified quiz mode in which there is some possibility of error because feedback is not immediate while original learning is taking place.

b) Pupils using the modified quiz mode will score higher on the criterion test than pupils using the test mode or the no-practice mode because there is less feedback in the latter programs and more possibility of error reinforcement.

c) Pupils using the no-practice mode will do better than those using the test mode which provides no feedback and may be accompanied by many error responses.

d) Pupils using the test mode will achieve the lowest mean score on the criterion test.

II There is an interaction effect between the presentation mode used and the intelligence of the pupil. It is therefore hypothesized that:

a) pupils who achieve high scores on scholastic aptitude tests will do just as well on programs which do not require overt responses as on programs which do require overt responses but which do not supply maximum feedback and reinforcement;

b) pupils who make low scores on scholastic aptitude tests will do better with programs requiring overt responses and supplying maximum reinforcement and feedback.

III The time taken to complete each program is a function of the amount of participation required of the pupil.

a) The no-practice mode will require the least amount of time because it requires reading only.

b) The quiz mode will be second because there is a delay after the presentation of each item while the machine is asked for the correct answer.

c) The modified quiz mode will rank third because time is needed to construct the answer and then to check it.

d) The test mode will take the most time because it is necessary to read the item until it is learned and then to think about the answer afterwards.

CHAPTER IV

METHOD AND EXPERIMENTAL DESIGN

Four methods were used to present programmed materials to the subjects in this study. The quiz and modified quiz programs were presented on the Didak 501, a partially adaptive teaching device. The no-practice group received their materials each day in the form of a mimeographed booklet. Insofar as wording permitted, the frames of this latter program were kept identical to those in the first two methods of presentation. The test mode group were also given their instructional materials in booklet form, but received each day an additional booklet containing a short quiz based on that day's lesson. The criterion test was bound in a separate booklet with the directions for taking the test affixed to the front of the booklet. A copy of the criterion test is to be found in Appendix B. Sample copies of the four programs have been reproduced in Appendix C.

The Program

The basic program consisted of 256 frames adapted from a programmed text called English 2600. The program dealt with grammatical analysis and synthesis involving simple, compound and complex sentences and the recognition of adjective and adverb clauses. (See Blumenthal, 1960, frames 692-965.)

This basic program was adapted to suit each instructional method and the material was presented as follows:

<u>Group Q (Quiz Mode)</u> The material was presented by means of the Didak 501. The pupil was instructed to read the item and to "think" the answer before pushing the lever revealing the correct answer for that frame. It was assumed that these instructions would produce a covert response during the interval between the presentation of the stimulus item and the correct response. This program was considered to be a modified form of the Skinner-type of constructed response, linear program.

<u>Group MQ (Modified Quiz Mode)</u> This group also used the Didak 501, but were asked to construct an answer to each frame before proceeding to the next. They then compared the written answer to the one provided by the machine. No correction was required before proceeding, but if the subject made a wrong response he was asked to depress the error counter using the device available on the machine. In this way the number of errors made each day by each pupil could be recorded. The pupil then proceeded to the next item.

<u>Group NP (No Practice Mode)</u> This group received the materials in booklet form. The information was printed in the form of frames identical to those used by groups Q and MQ, but the answers were provided as a part of the information given. In order to draw attention to these answers

they were underlined. This method was in some ways similar to a textbook type of presentation.

Group T (Test Mode) This group was given the material in textbook form. No underlining was used, but a quiz was given after each day's lesson. The quiz consisted of 10 items, approximately 50 per cent of which were later to be taken on the criterion test and 50 per cent of which were considered to be transfer items. The quiz was scored by the experimenter and returned the following day without discussion. This latter method may be considered to be similar to methods used in a conventional classroom.

Nature of the Test

The test used in this study was designed to serve two major purposes: a) as a pre-test for the purpose of choosing subjects for the experiment and b) as a criterion test to determine the amount gained by pupils used in the experiment. The actual unit of study surveyed by the test is called: "Understanding the Sentence Unit and Building Better Sentences". (Blumenthal, 1960)

The main objectives to be measured are set down in the test blueprint, Table III. These objectives are patterned after the six main educational objectives set down by Bloom *et al* (1956) in their compilation called a *Taxonomy of Educational Objectives*. Four of the six objectives from Bloom *et al* have been used, including knowledge, analysis, synthesis and evaluation.

The four subtopics in this unit are listed at the top of each column in Table II, while the seven objectives are listed in the left hand column. The percentages of the total number of items which were to be assigned to each separate topic and objective are listed below the respective headings. Each cell in Table III contains notes on the subject matter to be covered in the unit and below each note is the number of items which were used to test the content of that section of the work. The test contains 65 items for which a test time of 40 minutes was allotted. Of the total number of items, 42 are variations of the multiple choice question using either two, three or four alternatives. The remaining items require the pupils to construct answers. It was hoped that these items could be used to ascertain whether or not the pupils had adequately learned to use the grammatical knowledge of the unit in a practical situation. It was felt that only by this type of test, in which there was a wide variety of items, could the objectives and subject matter covered be adequately tested.

Nature of the Test Sample

The test was prepared for a Grade VII language class considered to be fairly typical of those found elsewhere in Alberta. It was administered to the entire Grade VII class of a school in Jasper Place. This involved a total of 92 subjects. The intelligence quotients of this group

TABLE II

TEST BLUEPRINT FOR UNDERSTANDING
THE SENTENCE UNIT AND BUILDING BETTER SENTENCES

EDUCATIONAL OBJECTIVES	A. SENTENCE PATTERNS	B. THE ADVERB CLAUSE	C. THE ADJECTIVE CLAUSE	D. SENTENCE FRAGMENT
1. Knowledge of Specific Facts (1.12) (5%)	Recalling Basic Sentence Components (2)*	Recalling Basic Components of Adverb Clauses (2)	Recalling Basic Components of Adjective Clauses	
2. Knowledge of Terminology (1.11) (15%)	Knowing Definitive Characteristics of Sentence Patterns (3)	Knowing Definitive Characteristics of Adverb Clauses (3)	Knowing Definitive Characteristics of Adjective Clauses (3)	
3. Knowledge of Conventions (1.21) (15%)	Knowing Characteristic Method of Presenting Sentences (3)	Knowing Characteristic Structure of Adverb Clauses (3)	Knowing Characteristic Structure of Adjective Clauses (4)	
4. Knowledge of Classifications and Categories (1.23) (20%)	Recognizing Basic Sentence Patterns (5)			Recognizing Sentence Fragments (3)
5. Analysis of Elements (4.10) (25%)	Recognizing Parts of Sentences (4)	Recognizing Adverb Clauses in Specific Sentences (8)	Recognizing Adjective Clauses in Specific Sentences (5)	
6. Production of a Unique Communication (5.10) (15%)	Building Good Compound and Complex Sentences (6)	Using Adverb Clauses to Improve Sentences (3)	Using Adjective Clauses to Improve Sentences (2)	
7. Judgment in Terms of Internal Evidence (6.10) (5%)	Recognizing Good Compound Sentences (4)			

* Numbers in brackets indicate the number of items in each category.

ranged from 75 to 149 on the Lorge Thorndike Verbal Intelligence Test. The age range of the group was from 12 years 6 months to 15 years 4 months. The test was also given to a heterogeneous class in a Calgary Public School, and this group was included in the calculation of the reliability coefficient.

Test Items

The test items were prepared by the writer and submitted to a group of school teachers to be checked for content and appropriateness. After the necessary revisions were made, the items were printed and set up in booklet form. The Kuder-Richardson Formula No. 21 was used to check the reliability of the test. This was calculated to be .91.

Although the test consisted of 65 items, it was hoped that the items would be sufficiently difficult so that even the brighter pupils might make substantial gains as a result of the instructional program. It was anticipated that the frequency distribution of the scores for the sample used on the pre-test would be positively skewed, discriminating more adequately at the higher score levels. Table III shows the frequency distribution of scores made by pupils on the pre-test.

Directions and Administration

The introductory page enclosed in each test booklet gave the testee the necessary instructions. Once the

TABLE XII

GROUPED FREQUENCY DISTRIBUTION OF SCORES
ON LANGUAGE PRE-TEST GIVEN TO 125 GRADE VII
STUDENTS FROM 2 ALBERTA SCHOOLS

Interval	Pupils Tested in Jasper Place School	Pupils Used in Exp. Study	Pupils Tested in Calgary School	Total
60 - 64	1			1
55 - 59	1		1	2
50 - 54	7		2	9
45 - 49	8		3	11
40 - 44	8		3	11
35 - 39	8		5	13
30 - 34	10	6	6	16
25 - 29	7	5	5	12
20 - 24	14	11	3	17
15 - 19	14	10	3	17
10 - 14	10	7	1	11
5 - 9	2	1	0	2
0 - 4	2		1	3
	N = 92	N = 40	N = 33	N = 125
Mean	29.1	20.75	32.4	30.04
S.D.	14.15	8.75	12.20	15.27
Median	26.7	20.2	32.0	29.5

background data had been completed by the subjects, the test administrator read the paragraph "Directions for Writing the Test" with the group. The command to begin the test was given; and the test period proceeded for forty minutes, after which time the booklets were collected.

A copy of the test, together with instructions for scoring, is presented in Appendix B. Each item was allotted one mark for a total possible score of 65. In order to correct for guessing the usual correction formula was used:

$$\text{Score} = \text{Rights} - \frac{\text{Wrongs}}{n-1}$$

where n equals the number of alternatives.

Table IV is a summary of the item difficulty and discrimination index for the total pre-test group where N is equal to 125.

<u>Subjects</u>

Of the three classes (92 students) tested in Jasper Place, one class had been directly taught some of the material. This was reflected in the scores on the pre-test. Therefore the one class was eliminated from the study. Of the remaining 60 pupils those with scores over 35 were also eliminated. Two pupils were excluded because there was reason to believe that their scores on the pre-test were invalid because of a language difficulty. Three more had extremely low intelligence scores and were also excluded. Thus 25 students remained. The middle 15 were discarded

TABLE IV

ITEM DIFFICULTY AND DISCRIMINATION INDEX OF 65 ITEMS BASED ON PRETEST SCORES (N = 125)

Item No.	p[1]	r[2]	Item No.	p	r	Item No.	p	r
1	.54	.64	21	.62	.68	46	.34	.48
2	.38	.28	22	.77	.95	47	.49	.45
3	.68	.37	23	.75	.89	48	.29	.59
4	.58	.45	24	.74	.70	49	.31	.54
5	.41	.28	25	.67	.35	50	.91	.70
6	.70	.68	26	.58	.70	51	.57	.61
7	.38	.51	27	.61	.62	52	.69	.31
8	.66	.73	28	.58	.86	53	.85	.29
9	.43	.61	29	.38	.65	54	.43	.68
10	.65	.34	30	.78	.57	55	.50	.15
11	.85	.22	31	.58	.65	56	.66	.30
12	.98	.05	32	.78	.34	57	.63	.44
13	.58	.41	33	.80	.70	58	.38	-.18
14	.94	.60	34	.38	.66	59	.80	.68
15	.73	.70	35	.42	.62	60	.37	.60
16	.72	.91	36	.64	.34	61	.64	.35
17	.72	.36	37	.51	.59	62	.57	.55
18	.72	.91	38	.36	.10	63	.69	.75
19	.75	.69	39	.58	.65	64	.59	.61
20	.79	.74	40	.86	.46	65	.50	.43
			41	.39	.58			
			42	.75	.48			
			43	.19	.31			
			44	.52	.76			
			45	.55	.75			

1 p = Item Difficulty, proportion right.

2 r = Discrimination Index from Tetrachoric Tables

to give two dichotomized groups of 20 high IQ and 20 low IQ students. The means of these two groups were now submitted to a t-test to determine whether or not they were significantly different from each other. A t of 7.72 was calculated indicating that with 38 degrees of freedom the difference was significant at the .001 level. The two groups were then subdivided into four groups of five each by using a table of random numbers.

Procedure

The pre-test was given one week prior to the beginning of the experiment. After the subjects had been assigned randomly to groups Q, MQ, MP and T, they were scheduled to work at the programs once a day for three successive days. The only instructions given consisted of telling the pupils that they were all participating in an experiment and that they would be tested immediately after the three-day learning period to see how much they had learned. No time limit was set for completion of each day's work and pupils were encouraged to work carefully. After some sample items had been attempted to familiarize each subject with the program and the machine, presentation of material was begun. A record was kept for each subject of time taken to complete each of the three sets of frames. In the case of group MQ each student also kept a record of the number of errors he had made. Members of group T were informed each day of the number of

errors made on the quiz for the previous day. They were allowed to check their papers, but discussion was not encouraged. The final test was given on the fourth day.

CHAPTER V

ANALYSIS OF THE DATA

The first step in the analysis of the experimental data was to submit this data to an analysis of variance and covariance. (See Winer, pp. 578-605.) The results of the analysis are presented in this chapter together with other pertinent data. Discussion of the experimental results is found in Chapter VI.

<u>Presentation Methods</u>

Preliminary to the analysis of variance and covariance, tests for the homogeneity of variances were conducted on the pre-test and post-test scores using Hartley's procedure reported in Winer, 1962, pp. 93-94. The results of these tests showed that the within-cells variances were homogeneous. In order to ascertain if it were appropriate to use a pooled within-class regression coefficient for the analysis of covariance, the test for homogeneity of within-class regression was conducted. (See Winer, p. 5[illegible].) The test for homogeneity was satisfied.

The summary of the analysis of variance and covariance is found in Table 7. The analysis of variance indicates a difference in presentation methods significant at the .01 level. When the covariate (pre-test score) is taken into account, the error variance is reduced and the difference among presentation methods reached an even higher level of significance at the .001 level.

A further examination of Table V indicates that there is no significant difference in IQ levels. An interaction effect is, however, observed and is significant at the .05 level. The estimate of the square of the within-cell correlation is .448. This indicates a correlation between pre-test and post-test scores of approximately .67. The mean square due to experimental error in the analysis of covariance is approximately .55, and this indicates the amount of error not predictable by the covariate.

Table VI shows the adjusted criterion means for cells, methods and IQ levels. This adjustment was made by the procedure reported in Winer, p. 599 using the common pooled within-class regression coefficient.

Table VII indicates the results of applying Duncan's Multiple Range Test (Edwards, 1960) to the adjusted treatment means of the criterion test. Inspection of Table VI indicates that the order of the magnitude of these means is Q, MQ, MP and T. There is no significant difference in mean achievement between groups Q and MQ or between groups T and MP. There is, however, a highly significant difference between these pairs of presentation methods--that is, mean achievement for groups Q and MQ is significantly greater than for groups T and MP.

TABLE V

TWO-WAY ANALYSIS OF VARIANCE OF POSTTEST SCORES

AT TWO LEVELS OF INTELLIGENCE AND

FOUR METHODS OF PRESENTATION

Source of Variation	SS	d.f.	MS	F
IQ Level	397	1	397	3.47
Methods	1510	3	503.3	5.03
Interaction	206	3	68.6	
Error	3207	32	100.2	
Total	5320			

$F_{.95}(3/32)$ 2.92 $F_{.95}(1/32)$ 4.17

$F_{.99}(3/32)$ 4.51 $F_{.99}(1/32)$ 7.56

TWO-WAY ANALYSIS OF COVARIANCE OF POSTTEST SCORES

AT TWO LEVELS OF INTELLIGENCE AND

FOUR METHODS OF PRESENTATION

Source of Variation	SS	d.f.	MS	F
IQ Level	222	1	222	4.01
Methods	2096	3	698.6	12.63
Interaction	430	3	173.3	3.13
Error	1769	31	55.3	
Total	4517			

$F_{.95}(3/31)$ 2.92 $F_{.95}(1/31)$ 4.17

$F_{.99}(3/31)$ 4.51 $F_{.99}(1/31)$ 7.56

TABLE VI

ADJUSTED CRITERION MEANS

IQ Level	Methods				
	Q	MQ	T	NP	All
A-High	40.1	40.2	27.6	26.8	33.7
B-Low	36.6	35.4	18.7	26.6	29.4
Both	38.2	37.8	23.1	26.7	31.5

TABLE VII

DUNCAN'S MULTIPLE RANGE TEST ON ADJUSTED MEANS OF PRESENTATION METHODS

	Methods				Shortest Significant Range	
	T	NP	MQ	Q	Sig. Level	
Mean	23.1	26.7	37.8	38.2	.05	.01
23.1		3.6	14.7	15.1	6.14	8.22
26.7			11.1	11.5	6.46	8.57
37.8				.4	6.67	8.72

Note: Any two treatment means not underscored by the same double line are significantly different.
Any two treatment means underscored by the same double line are not significantly different.

Intelligence Factors and Gain Scores

The analysis of the data up to this point has been based on the adjusted treatment means of the criterion test. In order to make the interpretation more meaningful gain scores were analyzed and related to intelligence test scores by correlation procedures. Table VIII summarizes the mean gain scores for the four experimental groups and also shows separate mean gains for each high and low IQ group. It will be seen, with reference to Table V, that there is no significant difference in achievement among IQ levels: however, an examination of the direction of the means indicates that in general, at least in this experiment, high IQ groups learn more efficiently than do low IQ groups when material is programmed.

Since such a relationship between IQ and gain scores is contrary to many research findings, Spearman's Rho was calculated between gain scores and intelligence for each of the four experimental groups. These correlation coefficients appear in Table IX and tend to indicate that in the case of the Quiz and Modified Quiz methods of presentation there is little correlation between intelligence and gain. In the Test and No Practice presentations there is a positive relationship between gain scores and IQ measures when the total group of ten is taken into account. Only the correlation for T is significant at the .05 level.

TABLE VIII

MEAN GAIN SCORES FOR FOUR METHODS OF PRESENTATION

AT TWO LEVELS OF IQ

IQ Level	Q	MQ	T	NP	Total Group
A-High	19.4	19.8	7.0	6.4	13.15
B-Low	16.4	15.0	-1.6	6.2	9.0
Both Levels	17.9	17.4	2.7	6.3	11.08

TABLE IX

CALCULATION OF SPEARMAN'S RHO BETWEEN IQ AND GAIN SCORES FOR FOUR PRESENTATION METHODS AT TWO LEVELS OF IQ

IQ Level	Method of Presentation			
	Q	MQ	T	NP
A-High	.05	.07	.37	.28
B-Low	-.33	.25	.12	.16
Both Levels	.12	.19	.59*	.45

*Significant at .05

Note: The pearson Product Moment Correlation Coefficient for total group (N=40) was calculated to be .18.

Time

Table X indicates the average time taken by each of the four experimental groups to complete the program. Group Q was able to complete the program most quickly with a mean time of 67.8 minutes. Groups T and MP did not differ markedly from this pattern, but Group MQ, who had to construct responses to each item, used considerably more time, ranging from 103-165 minutes, with a mean time of 130.3 minutes.

An attempt was made to use time as a second controlling variable in an analysis of covariance, but when the gain scores were plotted against the time scores, the regression lines so obtained were not parallel and it was concluded that a further analysis of covariance using time and gain scores was not justified.

Error Rate

It was possible to obtain error score during presentation only for group MQ. Table XI summarizes the data available showing: a) the mean and median error score for the ten subjects on each day's program; b) the percentage of frames incorrectly answered by subjects in MQ each day; and, c) the range of errors made on each day's program.

Spearman's rho was calculated between error score and time taken and was found to be .44. For an N of 10 this is not significant. The correlation between IQ and error rate was calculated to be .03, also not significant.

TABLE X

RANGE, MEAN AND MEDIAN TIME IN MINUTES

TO COMPLETE PROGRAMS

	Method of Presentation				Total Group
	Q	MQ	T	NP	
Mean	87.8	130.3	94.2	88.3	100.15
Median	86.0	127.5	96.0	92.5	98.0
Range	78-102	103-165	73-107	65-110	65-165

TABLE XI

ANALYSIS OF ERROR SCORES FOR SUBJECTS TAKING MQ PROGRAM

	Day		
	1	2	3
Range	2-16	6-22	7-23
Median	7.0	12.0	10.5
Mean	7.7	13.1	11.8
Percentage of Frames Incorrectly Answered	14	17	12

CHAPTER VI

DISCUSSION

The discussion which follows will be based primarily on the hypotheses proposed in Chapter III, p. 74. Additional findings will be discussed when these appear to be appropriate.

<u>Presentation Method</u>

Hypothesis I stated that the effectiveness of a particular type of presentation is determined by the interaction among a number of variables such as the mode of practice, the mode of response, reinforcement schedules, feedback factors and possibility of making error responses. Four methods of presentation were used in the experiment. It was hypothesized that the order of the magnitude of the criterion means would be as follows: Q, MQ, MP and T. An examination of the adjusted criterion means of Table VI, p. 91 reveals that the means actually did fall in this order. There was, however, no significant difference between methods Q and MQ or between methods MP and T.

The hypothesis noted above indicated that reasons for superiority of one method of presentation over another lay in the inherent qualities of that program. There may, however, be a number of other reasons for the superiority of methods Q and MQ. In the first place, the subjects in Q and MQ were using a machine for their work, while groups

NP and T were using booklets containing the programmed material. It is quite possible, therefore, that motivational effects generated by the machine could account for for some of the differences observed in learning. Further studies should be carried out in which all subjects either use a machine or have all of the material presented in booklet form

A second possible explanation for the superior achievement of groups Q and MQ may be stated in terms of arousal theory. It may be that the presence of the machine increases arousal sufficiently so that discrimination is more effective. This would be particularly true of those pupils known as underachievers who are usually unchallenged by ordinary classroom procedures. The opposite effect might account for the poor achievement of the test presentation group. If these pupils had fear of tests then the presentation of such a test would increase arousal beyond an optimal level and discrimination would be adversely affected.

A third possibility for the superior achievement of groups Q and MQ suggests itself. The subjects in these two groups worked individually while those in groups NP and T did their studying as a group. Thus, in addition to the way in which the material was presented, to tend to make these methods of presentation similar to conventional

methods of classroom presentation in the eyes of the pupils. Many psychological effects might arise if pupils sensed this similarity. For example, an instructional set based on past experiences with "language" as a subject might be present. Students who had failed to achieve in this area previously would certainly be conditioned to avoid activity related to any kind of language study. This conditioned inhibition might, however, be overcome if some motivational variable were altered--in this case, the introduction of the teaching machine.

Further weight is lent to the above argument by the fact that many of the pupils in group I actually made lower scores on the post-test than they did on the pre-test. Pupils are so conditioned toward the regular classroom test that they tend to make little use of them. In fact, the evidence in this experiment suggests that the tests may even confuse them in that they tend to learn the incorrect answer which they supply rather than the one supplied by the instructor when he corrects or "takes up" the test. This difficulty might be overcome if the correct answer were given immediately after the answer was constructed by the pupil.

In further research it may be necessary to use subject matter which is entirely unfamiliar to the pupil. At the same time we do need information about school subjects so

some of these problems cannot be avoided. There is the added possibility that programming may help to overcome some of the difficulties discussed above. For example, those pupils who dislike language and are not learning because of experience of failure or dislike of teachers could perhaps be re-trained by using programmed instruction.

Interaction

Hypothesis II, p. 75, stated that there is an interaction effect between presentation mode and IQ. It also indicated that low IQ pupils will learn more effectively when required to construct responses, but that there will be no differences for high IQ groups. Inspection of Table V, p. 90, indicates that there is a significant interaction between IQ and method; however, an inspection of the adjusted mean criterion scores indicates that the significance is due to a larger than expected difference for the test mode and a smaller than expected difference for the no-practice mode. Furthermore, since the quiz and modified quiz methods of presentation were superior for both high and low IQ groups, there is no useful interpretation of this interaction. It is possible that the dichotomy between low and high IQ groups was not sufficiently great for a clear-cut interaction to be observed. It is postulated, therefore, that a clear interaction would show if a much higher IQ group were used for the upper level in a similar experiment.

An examination of Table IX, p. 95, tends to support the hypothesis that low IQ pupils benefit from constructing responses. For group Q there is a negative relationship between IQ and gain. This correlation is not significant for an N of 5; but, taken together with the positive correlation of .25 for the low IQ group in MQ, it does suggest that constructed responses may be superior for pupils in the lower IQ categories.

Further examination of Table IX supports the hypothesis that correlations between gain and IQ scores are higher for classroom types of presentation. Pupils are used to this method of instruction and may therefore continue to work at their accustomed or conditioned levels. Subjects in groups Q and MQ, who used the machines, probably benefitted from the added motivational effects present and underachievers worked more nearly to capacity, thus reducing the correlation for the combined high and low IQ groups. Again caution must be observed in interpretation because of the small number of subjects in each cell. A change in rank in only one case may make a considerable difference in the rank correlation coefficient.

Time

Hypothesis III, p. 75, suggested that time taken to complete each program is a function of the amount of participation required of the pupil. It was postulated that the amount of time taken, from least to most, would follow

the order: MP, Q, MQ and T. An examination of Table XI, p.97, indicates that the actual order was: Q, MP, T and MQ. The hypothesis regarding time is therefore not supported by the experimental data. There was no significant difference in mean time taken among groups Q, T and MP, but group Q took significantly more time because they had to construct responses for each item.

It was mentioned previously that gain scores were plotted against time scores in order that time might be used as a second controlling variable in the analysis of covariance. The regression lines thus plotted were, however, not parallel and no analysis on this basis could be made. The regression lines for groups Q and MQ tended to be curvilinear. This suggests that there is probably an optimum time or pace for these programs and would argue for the use of machines that pace the program according to the ability of the individual pupil.

In the remaining programs there was a slight decrease in gain scores with time. It is possible that lack of motivation may account for this decrease. Again these were the programs which bore a strong resemblance to conventional types of classroom instruction. Since group MQ achieved the highest mean score on the criterion test and also took the most time, the instructor must decide whether or not the small advantage over group Q is worth the added time taken by group MQ. It would seem that for this type of material, at least, the Quiz mode is superior.

Error Rate

An examination of Table XI, p. 90, indicates that from 12-17 percent of the items were incorrectly answered by the subjects in MQ. This indicates that the program was too difficult for some of the pupils or that they tended to work carelessly. It will be observed also that error rate reached a maximum during the second day and then receded on the third. It is possible that as the novelty of the machine wore off the subjects began to work more carefullyand deliberately.

The correlation between error score and time taken was calculated to be .44. This could be interpreted to mean that those who take more time make more errors, but since the correlation is not significant any such interpretation should be tentative pending further investigation with larger experimental groups. In any case, it may again be a function of motivation. The correlation between IQ and error rate was found to be -.03. This indicates that for this group there was no relationship between IQ and the number of errors made.

CHAPTER VII

SUMMARY AND IMPLICATIONS

Much time is spent by teachers preparing pupils for examinations and in working to some criterion by employing various types of drills and reviews. The author of this study chose the topic of programmed instruction because he felt that in this new technique lay the answer to some pressing problems.

1. How can the schools provide for needed drill without the tremendous cost of teacher supervision? The teacher is most certainly necessary in the actual initial learning situation; but while the pupils are carrying out drill work the teacher, who is now, for the most part, a professional person, could be much more beneficially employed preparing new material and helping individual pupils.
2. How can the school provide for individual differences? It is true that administrative procedures have made some progress in taking care of individual differences, but much time is still devoted to the slower learner while the bright pupil remains unchallenged. If programmed instruction can be developed and used to the point where it can bring pupils up to a certain standard before they are allowed to go on to new work, the teacher can be assured that the class is where he wants to start rather than

finding that he must go back to where the least well informed pupil is.

3. How can we rid the teacher of the wearisome and rather unprofessional task of marking material that does not need expert attention? Teaching machines can certainly be developed that are capable of scoring objective types of items. This again would free the teacher for more professional activities.

After seeing the possibility of using programmed instruction to solve some of these problems it followed that some evidence was needed before any questions could be answered. Some of this evidence was available in the form of research literature; many questions, however, remained unanswered. This study was undertaken to answer at least some of these.

Summary of the Study

The basic assumption underlying this study was that it should be possible to develop a technology of instruction applicable to teaching of all kinds and for all purposes. It was realized, of course, that this would require much time and many studies, but the urgency of educational needs makes such a technology desirable and necessary. It was also assumed that many sub-varieties of instructional methods would be useful, and that these varieties would need to be submitted to individual test and research. The experimenter therefore developed a

language program which could be used to test the effectiveness of four methods of presentation and instruction. The relative effectiveness of these methods was postulated on the basis of the extent of the presence of certain variables which were an essential part of a particular method of presentation. Forty children - 20 high IQ and 20 low IQ - in a Grade VII class of language were chosen as subjects. The 20 pupils in each IQ level were assigned randomly to four groups of five each. It was hoped by this procedure that some information would be obtained about the interaction between intelligence and method of presentation. The subjects worked through the program on each of three successive days and were then given a test to see how much they had gained.

Analysis of the final test indicated clearly that programming based on the writing of separate frames, to which each student had to respond before presentation of another frame, was superior to text book types of presentation. Further analysis of the data indicated that though an interaction effect between intelligence and method of presentation was present, it had no useful interpretation.

Implications for Programming Early research reports on programmed instruction have given the impression that expensive hardware would be necessary for successful automation. This line of thinking no longer seems to be in vogue. The very fact that overt responses do not appear to be necessary for efficient learning would indicate that

simple devices will be sufficient for most of the tasks required of teaching machines. It is, after all, the program that is really central to the problem. This study has shown that sequence is important and that better results are obtained when these are set up in the form of individual frames to which the pupil must respond overtly or covertly.

The study has also shown that material such as language for junior high school can be effectively programmed. If this is true of language, it is probably also true of other subjects but further research in various subject areas is needed. At this time there is no reason to believe that we have reached a limit on the type of material that can be programmed. At the same time it must be admitted that very little is known about programming for the development of higher mental processes; but then, we know little about how to do this in the regular classroom. The only real limit to programming seems to be the ingenuity of the programmer. To program material for a specific subject, it might be most effective if a teacher, subject matter expert, psychologist and pupil worked together.

<u>Implications for Further Research</u> There are many other methods of presentation not tested in this study. It will be necessary for further research to develop more of these methods and to replicate this study in other subject areas using other grade levels and IQ levels.

One problem which faces the researcher is that of control. The analysis of covariance is a very useful technique, but there is no real guarantee that control of pre-experiment ability has actually been attained. In this study, for example, one of the basic factors in the experiment was the division of subjects into dichotomous groups on the basis of intelligence. Though the intelligence test used reports a high reliability, there still remains the possibility that the test does not measure maximum ability but rather the level at which the pupil is now functioning. This means that IQ scores, as any other achievement scores, may not be measuring accurately the potential of the pupil, especially the underachiever. Furthermore, the usual group intelligence test is heavily loaded with culturally biased items which may not be related to his ability to learn a specific subject at all. It might be beneficial for the researcher if he were to develop a pre-program which would give him an unbiased estimate of the ability of a pupil to perform in the particular area to be used. This would make grouping possible which would take into account the presence of the machine. The author is presently engaged in programming a group intelligence test in an effort to find out if motivational effects do differ when tests of maximum achievement are presented in this form.

The possibilities for further research cannot possibly all be outlined here. More knowledge is needed about the

writing of frames, about how to structure a certain subject, about the effects on pupils at various intelligence levels and about the whole process of learning in general. Some problems that arise are fairly simple and will be solved quickly. Others are difficult and involved and will take much time. If this study succeeds in laying the groundwork for further research it will have served a worthy purpose. Ramo (1957) has stated encouragingly: "Don't be ashamed of proposing a new idea. Though worthless today, in ten years it may be of no value whatsoever."

BIBLIOGRAPHY

Abel, L. B. "The Effects of Shift in Motivation Upon the Learning of a Sensory Motor Task." *Archives of Psychology*, No. 205, 1936.

Alter, M., Eigen, L., and King, S. "A Preverbal Number Concept Program." Center for Programmed Instruction New York, 1961.

Amsel, A. "Some Comments on the Manipulation of Error Responses and Reinforcement Schedules." Lowry AFB, Denver, Colorada. Unpublished Paper, September, 1957.

Barlow, J. A. "Project Tutor" *Psychological Reports* 6: 15-20, 1960.

Berlyne, D. E. *Conflict, Arousal and Curiosity*, McGraw Hill, 1960, New York.

Bloom, B. F. Editor: *A Taxonomy of Educational Objectives* Handbook I, Cognitive Domain, Longman's Green, New York, 1956.

Blumenthal, Joseph, C. *English 2600*, W. J. Gage Ltd., Toronto, 1960.

Blyth, J. W. "Teaching Machines and Human Beings." Paper presented at the 1959 annual meeting of American Council of Education, *The Educational Record*, 1960.

Briggs, L. J. "The Development and Appraisal of Special Procedures for Superior Students and an Analysis of the Effects of Knowledge of Results." *Abstracts of Doctor's Dissertations*, Ohio State University Press, 58: 41-49, 1949.

__________. "Teaching Machines," *USAF-NRC Symposium on Education and Training Media*. Washington, D. C.: National Academy of Science- National Research Council, 1960.

Bryan, G. L., and Rigney, J. W., "An Evaluation of a Method for Shipboard Training in Operations Knowledge." Department of Psychology, University of Southern California, Technical Report No. 18, September, 1956.

Carr, W. J. *Self-Instruction Devices: A Review of Current Concepts*. WPAFB, Ohio WADC Technical Report No. 59, August, 1959.

Cotterman, T. E. "Task Classification: An Approach to Partially Ordering Information on Human Learning." WADC Technical Note. 58-374, January, 1959.

Coulson, J.E., and Silberman, H. F. "Results of Initial Experiments in Automated Teaching." Santa Monica, Calif,.: Systems Development Corporation, Report No. SP 73, July, 1959.

Crowder, N. A. "Automatic Tutoring by Intrinsic Programing." Lumsdaine, A. A. and Glaser, R. (Eds.) Teaching Machines and Programmed Learning: A Source Book. Washington, D. C.: NEA, 1960, p. 286-298.

Edwards, A. L. Experimental Design in Psychology and Research. Rinehart and Co., New York, 1960

Eigen, L. D. and Komoski, P. K. "Research Summary No. 1 of the Collegiate School of Automated Teaching Project." New York, June 1960.

Eigen, L. D. and Filep, R. T. "A Comparison of Modes of Presenting a Programmed Instruction Sequence." The Center for Programmed Instruction, Inc., New York, December, 1961.

Evans J. L. Glaser, R. and Homme, L. E. "A Preliminary Investigation of Variation in the Properties of Verbal Learning Sequences of the Teaching Machine Type." Unpublished Paper, 1959.

__________, __________,________________, "The Ruleg System for the Construction of Learning Programs." Department of Psychology, University of Pennsylvania, 1960.

Falconer, G. A. "A Mechanical Device for Teaching Word Recognition to Young Deaf Children." Ed. D. Thesis, University of Illinois, 1959.

Ferster, C. D. and Sapon, S. M. "An Application of Recent Developments in Psychology to the Teaching of German." Harvard Educational Review, 5, No. 2: 451-467, 1958.

Finn, J. D. "Teaching Machines." NEA Journal 49: 41-44, 1960.

Fry, E. B. "A Study of Teaching Machine Response Modes." Lumsdaine, A. A. and Glaser, R. (Eds.) Teaching Machines and Programmed Learning: A Source Book. Washington, D. D.: NEA 1960, p. 469-474.

Gagne, R. M. and Dick, W. "Learning Measures in a Self-Instructional Program in Soving Equations." Department of Psychology, Princeton University, 1960.

Galanter, E. H. editor. Automatic Teaching: The State of the Art. New York: John Wiley and Sons, 1959.

Gotkin, L. G. and Goldstein, L. S. "Programmed Instruction for the Younger Learner: A Comparison of two Presentation Modes in two Environments." The Center for Programmed Instruction, New York, January, 1962.

Gilbert, T. F. "On the Relevance of Laboratory Investigation of Learning to Self-Instructional Programming." Lumsdaine, A. A. and Glaser, R.: (Eds.) Teaching Machines and Programmed Learning: A Source Book. Washington, D. C.: NEA, 1960, p. 475-485.

Green, Edward, J. The Learning Process and Programmed Instruction. Holt, Rinehart and Winston, Inc. New York, 1962.

Hilgard, E. R. "Learning Theory and its Applications." New Teaching Aids for the American Classroom, 1961.

Holland, J. G. "A Teaching Machine Program in Psychology." In Galanter, E. (Ed.) Automatic Teaching. New York, Wiley, 1959, (a) p. 69-82.

____________, "Teaching Psychology by a Teaching Machine Program." Unpublished Thesis, 1959 (b).

____________, and Porter, D. "The Influence of Repetition of Incorrectly Answered Items in a Teaching Machine Program. American Psychologist, 15: 423, 1960.

Ingles, J. H. "The Effectiveness of Programmed Instruction: Experimental Findings for 7070 Training." IBM Corporation, January, 1961.

Irion, A. L. "An Investigation of Four Modes of Operation and Three Types of Learning Tasks on the Improved Subject Matter Trainer." AFPTRC, Unpub. Draft, 1957.

Jersild, A. T. "Examination as an Aid to Learning." Journal of Educational Psychology, 20: 602- 609, 1929.

Jensen, B. T. "An Independent Study Laboratory Using Self-Scoring Tests." Journal of Educational Research, 43: 134-137, 1949.

Kahin, A. "Development and Evaluation of a Programmed Text in an Advanced Mathematical Topic for Intellectually Superior Fifth and Sixth Grade Pupils." Doctoral Dissertation Abstract, Florida State University, 1961.

Kaess, W. and Zeaman, D. "Positive and Negative Knowledge of Results on a Pressey-type Punchboard." Journal of Experimental Psychology, 60: 12-17, 1960.

Keislar, E. R. "Theoretical Aspects of Automated Learning." 1959 (a). Unpublished Paper read at APA.

_______________, "The Development of Understanding in Arithmetic by a Teaching Machine." Journal of Educational Psychology, 50: 247-253, 1959 (b).

Kendler, H. H. "Teaching Machines and Psychological Theory." Ch. XV in Galanter, E. (ed.) Automatic Teaching: The State of the Art. New York, Wiley and Sons, 1959.

Kennedy, R. L. "A Critical Review of a Portion of the Literature on Teaching Machines." Unpublished Paper. University of Alberta, April, 1961.

Little, J. K. "Results of Use of Machines for Testing and for Drill Upon Learning in Educational Psychology." Journal of Experimental Education, 3: 45-49, 1934.

Lumsdaine, A. A. and Glaser, R. Teaching Machines and Programmed Learning: A Source Book, Washington, D. C. NEA, 1960.

Meyer, Susan, R. "A Program in Elementary Arithmetic: Present and Future." Ch VI in Galanter, E. (Ed.) Automatic Teaching: The State of the Art, New York: John Wiley and Sons, 1959 (a), p. 83-84.

_______________, "Report on the First Run of a Junior High School Vocabulary Presented in Programmed Textbook." Unpublished. November 1959 (b).

Mayhew, D. J. and Johnson, A. F. "Teaching Machines." Journal of Engineering Education. 50: no. 1, 51-52, 1959.

Ordahl, Louise, E. and Ordahl, G. "Qualitative Differences Between Levels of Intelligence in Feeble-minded Children." Journal of Psycho-asthenics, Monograph Supplement 1915, 1, no. 2, 3-50.

Osler, Sonia, F. "Apparatus for Discrimination and Concept Formation Experiments, American Journal of Psychology, LXXIII, No. 3, 1960.

Pask, G. "Adaptive Teaching with Adaptive Machines." Lumsdaine, A. A. and Glaser, R. Teaching Machines and Programmed Learning: A Source Book. Washington, D. C. 1960.

Porter, D. "A Critical Review of a Portion of the Literature on Teaching Machines." Harvard Educational Review, 27: 126-147, 1957.

__________, "Some Effects of Year-long Teaching Machine Instruction." Ch VII in Galanter, E. (Ed.) Automatic Teaching: The State of the Art. New York: John Wiley and Sons, 1959, p. 85-90.

Pressey, S. L. "A Simple Apparatus which Gives Tests and Scores and Teaches." School and Society. 23: 373-376, 1926.

____________, "Some Items for a History of Automation of Teaching: Self Instructional Devices." APA, 1958.

Ramo, S. "A New Technique of Education." Engineering and Science Monthly, 21:1722, October, 1957.

Rath, G., Anderson, Nancy, S. and Brainerd, R. C. "The IBM Research Center Teaching Machine Project." Ch. XI in Galanter, E. (Ed.) Automatic Teaching: The State of the Art. New York: John Wiley and Sons, 1959, p. 117-130.

Robinson, F. G. "Some observations on the Research Literature on Programed Learning." A Talk Presented to the Canadian Teacher's Federation and Published in a Monograph entitles Programmed Learning and Its Future in Canada. 1961.

Roe, A et al, "Automated Teaching Programs Using Linear Programs." Automated Learning Research Project, Department of Engineering, University of California, Los Angeles, 1960.

Rothkopf, E. Z. "Some Research Problems in the Design of Materials and Devices for Automated Teaching." Lumsdaine, A. and Glaser, R. (Eds.) Teaching Machines and Programmed Learning: A Source Book, Washington, D. C. NEA, 1960, p. 318-328.

Rushton, E. W. "Roanoke Experiment," Nations Schools, February, 1961, p. 76-79.

Shafer, S. H. "Teaching Machines and the Social Studies." Social Education, February 1961, p. 85-86.

Silberman, H. F. "Introductory Description of Teaching Machines." Unpublished Paper, 1959.

______________ and Coulson, J. "A Draft Summary of Findings in an Exploratory Teaching Machine Study." Automated Teaching Bulletin, 1: 16-18; 2: 35-37, December, 1959.

Skinner, B. F.

Skinner, B. F. "Scince of Learning and the Art of Teaching." <u>Harvard E. Rev.</u>, 24, 86-97, 1954.

Skinner, B. F. <u>Verbal Behavior</u>, Appleton-Century-Crofts, New York, 1957.

______________, "Programming Verbal Knowledge for Automated Instruction in Schools and Colleges." <u>American Psychologist</u>, 1958.

______________ and Holland, J. G. "The Use of Teaching Machines in College Instruction." Final Report to the Fund for Advanced Education, August, 1958, p. 159-172.

Stephens, A. L. "Certain Special Factors Involved in the Law of Effect." <u>Abstracts of Doctor's Dissertations</u>, Ohio State University, 64: 89-93, 1953.

Stolurow, L. M. <u>Teaching by Machine</u>. U. S. Department of Health, Education and Welfare, Washington, 1961.

Taylor, Janet, A. and Spence, K. W. "The Relationship of Anxiety Level to Performance in Serial Learning." <u>Journal of Experimental Psychology</u>, 44: 61-64, 1952.

Walker, C. C. and Stolurow, L. M. "A Comparison of Overt and Covert Response in Programmed Learning." <u>The Journal of Educational Research</u>, 55: 9, July, 1962.

Winer, B. J. <u>Statistical Principles in Experimental Design</u>. McGraw-Hill Book Co., Inc., New York, 1962.

Woodrow, H. "The Effect of Practice on Groups of Different Initial Ability." <u>Journal of Educational Psychology</u>, 29: 268-276, 1938.

Zeaman, D. "Skinnerian Theory of Teaching Machines." In Galanter, E. (Ed.) <u>Automatic Teaching: The State of the Art</u>. John Wiley and Sons, New York, 1959, p. 167-176.

APPENDIX A

TABLE I

AGE, LORGE-THORNDIKE VERBAL IQ's, PRE-TEST AND POST-TEST FOR PUPILS USED IN JASPER PLACE EXPERIMENT N = 40

Pupil No.	Age	L. Th. IQ	Pre-Test	Post Test
Group Q				
1.	12-8	106	30	47
2.	12-7	106	15	44
3.	12-6	122	34	53
4.	13-7	115	12	31
5.	12-8	110	16	29
6.	13-7	101	23	42
7.	13-11	97	14	30
8.	12-11	105	26	39
9.	13-3	82	11	29
10.	14-1	98	11	27
Group MQ				
11.	13-8	110	19	40
12.	13-6	113	15	28
13.	12-6	112	25	46
14.	13-6	112	20	38
15.	13-6	111	20	46
16.	13-7	101	19	22
17.	13-9	100	20	32
18.	12-10	103	23	48
19.	13-10	90	22	38
20.	12-6	94	23	42

Pupil No.	Age	L. Th. IQ	Pre-Test	Post-Test
Group T				
21.	13-11	111	18	22
22.	12-10	119	28	29
23.	12-11	106	18	35
24.	12-10	112	30	53
25.	12-10	106	24	14
26.	13-10	99	24	26
27.	13-1	103	27	21
28.	13-10	95	14	7
29.	14-4	80	10	23
30.	14-0	80	20	10
Group NP				
31.	14-6	116	30	33
32.	13-2	114	30	36
33.	12-10	108	14	21
34.	12-2	106	17	30
35.	12-4	104	15	18
36.	13-0	103	32	46
37.	13-10	99	26	27
38.	13-7	98	20	25
39.	13-2	103	15	15
40.	12-6	97	7	18

TABLE II

IQ, PRETEST, AND POSTTEST
MEANS AND STANDARD DEVIATIONS

IQ Level	Test	Presentation Method								All Methods	
		Q		MQ		T		NP			
		$\bar{X}$	S	$\bar{X}$	S	$\bar{X}$	S	$\bar{X}$	S	$\bar{X}$	S
A-High	IQ	111.8	6.1	111.6	1.1	110.8	4.9	109.6	4.6	110.95	4.63
	Pretest	21.1	9.5	19.8	3.2	23.6	4.9	21.2	7.3	21.5	6.57
	Posttest	40.8	9.2	39.6	6.6	30.6	13.2	27.6	7.1	34.6	10.95
B-Low	IQ	96.5	9.0	98.0	5.3	91.4	9.6	98.6	4.8	96.5	7.57
	Pretest	17.0	6.3	21.4	1.9	19.0	6.3	20.0	8.7	19.4	8.27
	Posttest	33.4	6.0	36.4	8.9	17.4	7.5	26.2	10.8	28.4	11.22
Both Levels	IQ	104.2	12.1	104.8	7.8	101.1	12.3	104.1	7.2	103.7	9.52
	Pretest	19.2	8.0	20.6	2.7	21.3	6.1	20.6	8.0	20.4	6.69
	Posttest	37.1	8.8	38.0	8.0	24.0	12.6	26.9	9.1	31.5	18.81

APPENDIX B

NAME______________________________, __________________________________ CLASS ____________________________

Last Name First Name

DATE OF BIRTH ________, _____________, ________ SCHOOL____

Year Month Day

WHAT SCHOOL DID YOU ATTEND LAST YEAR? ____

WHERE IS THIS SCHOOL? ____

DIRECTIONS FOR WRITING THE TEST

This language test consists of eight parts each of which contains a number of questions. You are to read the instructions for each part and then to do the questions that follow by placing your answers in the spaces provided in your test booklet. The marks you get on this paper will not go on your report cards. Some of the questions are hard and I do not expect you to be able to answer all of them, but I want you to do the best you can. Marks will be taken off for guessing, so do not guess unless you are fairly sure that you know the right answer. You have forty minutes to finish the test. When you finish turn the paper over and go on quietly with some other work.

DO NOT TURN THE PAGE UNTIL I SAY, "BEGIN."

LANGUAGE TEST

READ THE INSTRUCTIONS FOR EACH PART OF THE TEST CAREFULLY AND BE SURE THAT YOU PLACE YOUR ANSWER IN THE PLACE PROVIDED.

I. Groups of words may be classified or labelled as:

A. A simple sentence
B. A compound sentence
C. A complex sentence
D. A fragment

Each of the following groups of words may be classed as one of the above types. Indicate your choice by placing the capital letter corresponding to your choice in the blank in the right-hand margin.

1. Tom is a boy who has many interests. _____

2. I went to the store while my brother went to the bank. _____

3. During the Easter holidays, when there is no school. _____

4. He struck out, and the pitcher cheered. _____

5. Which one of the boys who was in the building when the fire started? _____

6. Although the bell had rung, Harry was in no hurry to get to school. _____

7. I have a cat, a dog and two rabbits in my very own zoo. _____

8. Sam does well in arithmetic and Spanish, but he does poor work in social studies. _____

9. He saw a tree that had been blown over during the hurricane. _____

10. Into the lonely forest Harry plunged. _____

II. One of each of the following pairs of compound sentences is good because it combines similar ideas of equal importance. The other is poor because it combines unlike ideas. Indicate which is the good sentence in each pair by placing the appropriate capital letter which corresponds to your choice in the space at the right.

A. The dress was a bargain, but I had spent all my money.
B. The dress was a bargain, and my friend had one like it. _____

A. Ken needed money for college, and his uncle owned a gas station.
B. Ken needed money for college, and his uncle offered him a job. _____

A. We had turkey for dinner, and I prefer the dark meat.
B. Dad prefers the light meat, and I prefer the dark meat. _____

A. Sally types with great speed, and her work is accurate.
B. Sally types with great speed, and she has a new typewriter. _____

III. If the underlined group of words is an Adjective Clause, write Adj. in the space on the right. If the underlined group of words is an Adverb Clause, write Adv. in the space on the right.

15. When a union and a company cannot agree, a strike may occur. ______

16. We have a number of customers who always phone their orders. ______

17. I prayed for snow so that I could try out my new skis. ______

18. Vic has a new watch which tells the day of the month. ______

19. The lady whose recipe won first prize was not in the audience. ______

20. I always read his article, although I frequently disagree with him. ______

21. Although he was a very old man, he still seemed to have a lot of energy. ______

22. The watch, which John had given me, stopped. ______

23. The cowboys rode into town after the sun went down. ______

24. When dinner is ready, I shall call you. ______

IV. Eliminate the "and" by changing the underlined statement to the kind of word group indicated in the parentheses. Rewrite each sentence.

25. We have had little rain and the crops will be poor (adverb clause)

____

____

26. I got the advice of a friend, and he knows a lot about cars. (adjective clause)

____

____

27. John never gains weight, and he is a big eater. (adverb clause)

____

____

28. We gave the bike to a child, and his father is out of work. (adjective clause)

____

____

29. I took the clock apart, and I couldn't put it back together again. (adverb clause)

____

____

V. Place the letter which corresponds to the best answer in the blank space on the right.

EXAMPLE: The planet which is closest to the sun is
A. Venus C. Mercury
B. Pluto D. Jupiter ___C___

30. Which of the following is NOT a conjunction?

A. and C. but
B. however D. or ______

31. Which one of the following words is an adverb clause signal?

A. but C. for
B. when D. none of these ______

32. Here is a sentence with a word missing:

"__________Sandy eats heartily, he cannot gain weight."

Which adverb clause signal below would bring out the meaning most clearly?

A. because C. although
B. as if D. whenever

33. Which one of the following sentences is NOT a complete sentence?

A. Riding down the road in the afternoon while the sun bakes the earth roundabout.
B. John was riding along the road on a hot summer afternoon.
C. On a hot summer afternoon the sun baked the earth. ______

34. A compound sentence consists of

A. A compound subject and a compound predicate.
B. Two subjects and one predicate.
C. One subject and two predicates.
D. Two subjects and two predicates. ______

35. "Some of my friends and neighbors came to supper, and we sang and played all evening."

Which of the following statements is true of the sentence above?

A. The first "and" connects two parts of a compound sentence.
B. The second "and" connects two parts of a compound sentence.
C. The third "and" connects two parts of a compound sentence.
D. This is not a compound sentence. ______

36. Which pair of sentences would make the best compound sentence when combined?

A. The dog was biting the man. He did not like it.
B. She has long curls. Sometimes she plays baseball.
C. Carl is a strange boy. The teacher likes all of the children in the class.

7. Which of the following groups of words is an adverb clause?

A. the house is burning
B. if the house is burning
C. burning in the house
D. the burning house

8. A complex sentence must contain

A. a clause
B. a phrase
C. a compound subject
D. two clauses

9. "While flying to Europe, Mr. Smith collided with an albatross."
The underlined group of words is

A. a prepositional phrase
B. an adjective clause
C. a word group
D. none of these

0. Which one of the following words could NOT be used to begin an adjective clause?

A. who
B. which
C. that
D. yet

1. An adjective clause

A. does the work of a simple subject
B. does the work of a compound subject
C. does the work of an adverb
D. none of these

2. "The man robbed the bank, who went into that hotel."
In this sentence the underlined adjective clause modifies

A. robbed
B. bank
C. man
D. who

3. We can tighten up a weak compound sentence by

A. using a compound subject
B. replacing part of the sentence with a preposition
C. using a different conjunection
D. changing one of the statements to an adjective clause

4. "The man in the black suit is a secret agent."
The underlined words form

A. an adjective clause
B. a prepositional phrase
C. a noun clause
D. an adverb clause

VI. Change the following compound sentences to complete sentences by crossing out the conjunction and placing the correct clause signal to replace it in the parentheses at the end of the sentence.

45. John can play football, and he once had polio. ( )

46. I went to the store, but John went to the park. ( )

47. Mother was knitting a scarf and the roast was buring. ( )

48. The bridge had collapsed, and the train had to detour. ( )

49. Ernie had a sore foot and it kept him from running in the race. ( )

Vii. Underline the best answer in each of the following.

50. A (simple, compound) sentence can be divided into two separate sentences.

51. We have been to Seattle and learned a lot about the World Fair. Should a comma be inserted after "Seattle?" (Yes, No).

52. We have been to Seattle, and we learned a lot about the World Fair. "We" has been added to the sentence in number 51. The comma after "Seattle" is now (right, wrong).

53. The piano played, the violin sang. In this sentence the (subject, conjunction) is missing.

54. If we wish to contrast two ideas, we should use the conjunction (and, but).

55. A clause is a group of words which has a subject and a verb.)It does, does not) make sense by itself.

56. An adjective clause does the work of a single (adjective, conjunction.)

57. Mary passed the examination although she did not study for it. Because this sentence contains a clause, it is a (compound, complex) sentence.

58. Adjective clauses are just as movable as adverb clauses. (true, false).

59. When we wish to refer to a person, we should use (which, who).

60. Adjective clause signal are (adjective, pronouns).

61. If we wish to change a compound sentence to a complex sentence, we should (add, remove) the conjunction "and".

62. An adjective clause usually comes (before, after) the noun or pronoun it modifies.

63. A clause which tells when something happened is most likely an (adjective, adverb) clause.

64. A complex sentence (can, cannot) be divided into two separate sentences.

65. A clause signal can never consist of more than one word (true, false).

SCORING PROCEDURE FOR CRITERION TEST

Section I - Items 1-10 - Score = Right minus $\frac{\text{Wrong}}{n - 1}$

(where n = number of alternative choices in each item)

Section II - Items 11-14 - Score = Right minus Wrong

Section III - Items 15-24 - Score = Right minus Wrong

Section IV - Items 25-29 - Score = Number of Items Correct

Section V - Items 30-44 - Score = Right minus $\frac{\text{Wrong}}{n - 1}$

Section VI - Items 45-49 - Score = Number of Items Correct

Section VII - Items 50-65 - Score = Right minus $\frac{\text{Wrong}}{n - 1}$

APPENDIX C

Samples of Programs Used in Jasper Place Experiment

Glen / buys old cars.

This sentence can be divided into two major parts: the complete subject and the complete pred______.

Glen and Chris / buy old cars.

predicate

Although the complete subject now contains two simple subjects, the sentence can still be divided into ______ major parts. (How many?)

Glen and Chris / buy old cars and rebuild them

two

Now the complete subject has two parts connected by the word "and," and the complete predicate has two parts also connected by the word ________.

Glen and Chris/buy old cars and rebuild them.

and

Because the complete subject has more then one part, we say the subject is compound. Because the complete predicate also has more than one part, we say that the predicate is ______.

Glen and Chris / buy old cars and rebuild them.

compound

Although both the subject and the predicate are compound, we can still divide the sentence into two parts: the complete subject and the complete __________.

A sentence that can be divided into two parts--a subject and a predicate--is a simple sentence. A simple sentence may have a compound subject and/or a compound __________.

predicate

My mother is English. My father is French.

predicate

Each of the above sentences can be divided into a subject and a predicate. Therefore each of the above sentences is a (simple, compound) sentence.

My mother is English and my father is French.

simple

The two sentences are now joined into a single sentence by the conjunction or connecting word ______.

My mother is English and my father is French. — and

This is not a simple sentence because each of its two parts now has its own subject and ______.

Our school is small, but we have good teams. — predicate

This sentence was made by joining two simple sentences with the conjunction ________.

You must shut the gate, or the dog will get out. — but

This sentence was made by joining two simple sentences with the conjunction ________.

A sentence made by joining two (or more) simple sentences with the conjunction and, but or or is called a compound sentence. — or
The most common conjunctions that connect the two parts of a compound sentence are and, but and ____.

___ __________ (, and
(, but ___ __________
(, or — or

Notice that in a compound sentence there is a subject and a predicate both before and after the __________ or connecting word.

conjunction

A simple sentence cannot be divided into two separate sentences.
A ________ sentence can be divided into twoe separate sentences.

compound

a. Our class made birdhouses and sold them.
b. Our class made birdhouses and the church soldthem.

Which sentence has both a subject and a predicate after the conjunction "and"? (a or b?)

b

Glen / buys old cars.

This sentence can be divided into two major parts: the complete subject and the complete predicate.

Glen and Chris / buy old cars.

Although the complete subject now contains two simple subjects, the sentence can still be divided into two major parts.

Glen and Chris / buy old cars and rebuild them.

Now the complete subject has two parts connected by the word "and," and the complete predicate has two parts also connected by the word "and".

Glen and Chris / buy old cars and rebuild them.

Because the complete subject has more than one part, we say the subject is compound.
Because the complete predicate also has more than one part, we say that the predicate is compound.

Glen and Chris / buy old cars and rebuild them.

Although both the subject and the predicate are compound, we can still divide the sentence into two parts; the complete subject and the complete predicate.

A sentence that can be divided into two parts--a subject and a predicate--is a simple sentence.
A simple sentence may have a compound subject and/or a compound predicate.

My mother is English. My father is French.

Each of the above sentences can be divided into a subject and a predicate.
Therefore each of the above sentence is a simple sentence.

My mother is English, and my father is French.

The two sentences are now joined into a single sentence by the conjunction, or connecting word, and.

My mother is English, and my father is French.

This is not a simple sentence because each of its two parts now has its own subject and predicate.

Our school is small, but we have good teams.

This sentence was made by joining two simple sentences with the conjunction but.

You must shut the gate, or the dog will get out.

This sentence was made by joining two simple sentences with the conjunction or .

A sentence made by joining two (or more) simple sentences with the conjunction and, but, and or is called a compound sentence. The most common conjunctions that connect the two parts of a compound sentence are and, but, and or .

___ ________________ (, and
(, but ______ ________
(, or

Notice that in a compound sentence there is a subject and a predicate both before and after the conjunction or connecting word.

A simple sentence cannot be divided into two separate sentences. A compound sentence can be divided into two separate sentences.

a. Our class made birdhouses and sold them.

b. Our class made birdhouses and the church sold them.

Sentence b has both a subject and a predicate after the conjunction "and."

a. Our class made birdhouses and sold them.

b. Our class made birdhouses and the church sold them.

Sentence b can be divided into two separate sentences.

a. Our class made birdhouses and sold them.

b. Our class made birdhouses and the church sold them.

Sentence b is a compound sentence.

Our class made birdhouses and sold them.

This is not a compound sentence because there is not a subject and predicate, but only a predicate after the conjunction "and."

UNIT I

A sentence can be divided into two major parts: the complete subject and the complete predicate.

Glen / buys old cars.

In this sentence there are two major parts, the complete subject "Glen" and the complete predicate "buys old cars."

Glen and Chris / buy old cars.

Although the sentence now contains two simple subjects, the sentence can still be divided into two major parts.

Glen and Chris / buy old cars and rebuild them.

Now the complete subject has two parts connected by the word "and", and we say that the subject is compound. The complete predicate consists of two parts and so the predicate, too, is said to be compound. Now a simple sentence can be divided into two parts, a subject and a predicate. If either the subject or the predicate are compound, the sentence is still simple.

My mother is English. My father is French.

Each of these two sentences can be divided into a subject and a predicate. Therefore each sentence is a simple sentence.

My mother is English and my father is French.

The two sentences have now been combined into a single sentence by the conjunction, or connecting word "and". Each of the two parts of the sentence now has its own subject and predicate. The sentence is no longer a simple sentence; it is now a compound sentence.

Our school is small, but we have good teams.

This compound sentence was made by joining two simple sentences with the conjunction "but".

You must shut the door, or the dog will get out.

This sentence was made by joining two simple sentences with the conjuntion "or". A sentence made by joining two (or more) simple sentences with the conjunctions and, but or or is called a compound sentence. These are the three most common conjunctions.

(, and
—— ———— (, but —— ————
(, or

Notice that in a compound sentence there is a subjct and a predicate both before and after the conjunction or connecting word. Remeber also that a simple sentence cannot be divided into two separate sentences, but that a compound subject can.

Our class made birdhouses and sold them.

This is not a compound sentence because there is not a subject and predicate but only a predicate after the conjunction "and". In this sentence the conjunction merely connects two predicates.

Mr. Rice excused himself and he and his wife left the meeting.

There are two "and's" in this sentence. The first "and" connects the two parts of this compound sentence.

Several boys and girls wrote and produced this program.

This sentence also contains two "and's." We notice, though, that neither one of these is used to connect the two parts of a compound sentence.

Here are some of the things that we know about simple and complex sentences.

1. A simple sentence cannot be divided into two separate sentences.
2. A compound sentence can be made by joining two simple sentences with a conjunction.
3. A compound sentence can be divided into two separate sentences.
4. In a compound sentence, there is a subject and a predicate both before and after the conjunction.
5. The three most common conjunctions are and, but and or.

Only similar ideas of equal importance should be combined into a compound sentence.

a. The lightning flashed, and the thunder rumbled.
b. The lightning flashed, and I was walking with Carl.

In sentence a. the ideas fit together best.

The lightning flashed, and the thunder rumbled.

Because both parts of this sentence are about weather conditions, this compound sentence is good.

One of us washes the dishes, and the other dries them.

This is a good compound sentence because both parts are about the washing of dishes. Now look at this sentence:

The food in our cafeteria is good. The prices are reasonable.

Note that the subject of the first sentence is "food"; the subject of the second sentence is "prices." Although each sentence has a different subject, both sentences are about the cafeteria's food--the first about its quality and the second about its price. Both of these sentences are about food, and the ideas are of equal importance.

Lois has beautiful eyes, and her hair is most attractive.

Both statements in this compound sentence are about Lois's appearance and are of equal importance.

Which of the following sentences is the better compound sentence?

a. Dad rented the house on Monday, and we moved in on Saturday.
b. Dad rented the house on Monday, and it has three bedrooms.

Because the two ideas in sentence a. are more alike, it is a better sentence.

Don't write two sentences together unless you use a conjunction to connect them.

The lightning flashed, the thunder rumbled.

This sentence is wrong because there is no conjunction to hold the two statements together.

a. I rang the doorbell, a child came to the door.
b. I rang the doorbell, and a child came to the door.

Sentence a. is wrong because the conjunction is missing.

Don't use "and" where "but" would bring out the contrast or difference between two things more clearly

We asked his permission, but he refused.

In this sentence "but" brings out the meaning most clearly. Use the conjunction that brings out the meaning most clearly. Notice the sentences below.

a. We asked his permission and he consented.
b. There was a stop sign at the corner, but Dick didn't see it.
c. The sign was too small, and many people failed to see it.